Annex to:

EFSA NDA Panel, 2022. Scientific Opinion on the Tolerable Upper Intake Level for dietary sugars. EFSA Journal 2022;20(2):7074. doi:10.2903/j.efsa.2022.7074

Annex N – Reanalysis of individual data on sugars intake and dental caries in children as preparatory work for the setting of a Tolerable Upper Intake Level for dietary sugars

Summary

EFSA requested all the authors of the observational studies potentially eligible for the assessment of the association between sugars intake and dental caries share individual data. The investigators of two prospective cohort studies (Iowa Fluoride Study, STRIP-2 study) on sugars intake and dental caries in children provided the data sets for the current analyses.

To address both excess zeros (i.e. a large number of zero counts) and overdispersion (i.e. excess variation relative to the Poisson distribution) that characterise the distribution of caries counts, zero-inflated negative binomial (ZINB) regression models were considered in addition to the more traditional approaches [Poisson, negative binomial, zero-inflated Poisson (ZIP) regression], to test the associations of sugars intake with caries increments while taking into account the available potential confounders. Dental caries was also analysed as indicators (i.e. occurrence/non-occurrence of new caries on tooth surfaces/teeth) as an alternative (and simplified) approach to count modelling. Logistic regression was applied to model the odds of occurrence in relation to sugars intake and to test the role and impact of potential confounders.

The original data set of the Iowa Fluoride Study cohort included 198 subjects (aged 5–9 years old), of which data from 192 subjects were used in the re-analyses. The association of total sugars intake with the outcome 'any surfaces with new non-cavitated or cavitated caries or filling between ages 5 and 9' was tested first in univariable and then in multivariable models; in the final model, which included age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old and tooth-brushing frequency as potential confounders, the adjusted odds ratio (OR) per 10 g/day increase in total sugars intake was 0.93 (95% confidence interval (CI): 0.83, 1.04). The same modelling strategy applied to the outcome 'Count of surfaces with new non-cavitated or cavitated caries or fillings between ages 5 and 9' did not show different results, as the adjusted incidence rate ratios (IRR) per 10 g/day increase was 0.97 (95% CI: 0.91, 1.04). In the final model that tested for the association of sugar-sweetened beverages intake and the same indicator outcome, the adjusted OR per 100 mL/day increase in intake was 1.01 (95% CI: 0.85, 1.21); similarly, there was no evidence for an association with the count outcome, with an adjusted IRR per 100 mL/day increase in intake of 1.00 (95% CI: 0.90, 1.10). The associations of fruit juices intake with the indicator outcome (adjusted OR per 100 mL/day increase was 0.83; 95% CI: 0.55, 1.26) and the count outcome (adjusted IRR per 100 mL/day increase was 0.96; 95% CI: 0.75, 1.24) were tested in similar multivariable models as for the other two exposures.

The original data set of the STRIP-2 study includes 148 subjects, of which 128 were used in the 3–6-year-old children analyses and 81 in the 12–16-year-old children analyses. The association of sucrose intake at age 3 years with the outcome 'any new d3mft between ages 3 and 6 years' was tested first in univariable and then in multivariable models; in the final model, which included daily tooth-brushing (Y/N) as potential confounder, the adjusted OR per 10 g/day increase in sucrose intake was 1.64 (95% CI: 1.13, 2.37). The estimate of the adjusted IRR per 10 g/day increase in intake from a modelling strategy applied to the outcome 'counts of new d3mft between ages 3 and 6 years' that included also caries experience at 3 years of age was 1.21, but it did not reach statistical significance (95% CI: 0.91, 1.61). The final model for the association of sucrose intake at age 12 years with the outcome 'any new D3MFT between ages 12 to 16' included intervention group, caries experience at 12 years of age daily

tooth-brushing (Y/N) at 12 years old as potential confounders; the adjusted OR per 10 g/day increase in sucrose intake was 0.95 (95% CI: 0.68, 1.34); similarly, there was no evidence for an association with the corresponding count outcome measured between ages 12 and 16, with an adjusted IRR per 100 mL/day increase in intake of 0.99 (95% CI: 0.84, 1.18).

Table of Contents

1. Methods

1.1. Data

EFSA requested all the authors of the observational studies potentially eligible for the assessment to share individual data. The current analyses are based on the data sets provided by the investigators of two prospective studies on sugars intake and caries in children (see Sections 10.1 and 10.2.2. in the Opinion). These are the Iowa Fluoride Study (IFS) cohort (Chankanka et al., 2011) and the STRIP-2 study (Karjalainen et al., 2001; Karjalainen et al., 2015).

A detailed description of the original published studies can be found in the evidence table for observational studies on dental caries (**Appendix M** to the Opinion). In the current report a description of the two data sets is included, and the re-analyses reported separately, as the two cohorts measured different exposures and endpoints outcomes (no pooling possible).

1.2. Statistical analysis

1.2.1. Caries data as counts

Caries counts are not generally approximated by a normal distribution; a Poisson distribution is usually considered appropriate for count data, but caries tends to exhibit overdispersion (i.e. excess variation relative to the Poisson distribution), possibly due to dependency of caries surfaces within an individual. The negative binomial distribution can be more appropriate for dental caries indices where overdispersion occurs (i.e. when the conditional variance exceeds the conditional mean).

Furthermore, distributions of caries counts are increasingly characterised by a large number of zero counts, with proportions in excess of what is expected under the Poisson and negative binomial distributions. To handle such 'excess zeros' the zero-inflated Poisson (ZIP) regression for modelling the decayed, missing, and filled teeth or surfaces indices (dmft; dmfs; DMFT; DMFS) has been proposed (Preisser et al., 2012). Two kinds of zeros are thought to exist in the data, 'true zeros' and 'excess zeros'. The latter are assumed to be generated by a separate process from the count values and modelled independently. Zero-inflated models estimate two equations simultaneously, one for the count model and one for the excess zeros (logit model).

ZIP models account for large counts of zeros, but they do not adequately account for data that have sizeable numbers of large caries counts. To address both excess zeros and overdispersion, zero-inflated negative binomial (ZINB) regression models have been proposed (Lewsey and Thomson, 2004) and were considered in the current analyses, in addition to the more traditional approaches, to test the associations of sugars intake with caries increments while taking into account the available potential confounders.

1.2.2. Caries data as indicators

As in the original analysis of one of the two data sets, dental caries was also analysed as indicators (i.e. occurrence/non-occurrence of new caries on tooth surfaces/teeth) as an alternative (and simplified) approach to count modelling. Logistic regression was applied to model the odds of occurrence in relation to sugars intake and to test the role and impact of potential confounders.

All data were analysed using Stata 15.1 (Scott Long and Freese, 2014); statistical significance was considered at the 5% level, but confidence interval (CI) estimation was used as the basis to interpret the results of the analyses.

2. Iowa Fluoride Study cohort

2.1. Data set description

Variable category	Description
	Subject identification number (198 subjects)
Intakes of total sugars, sugar-sweetened beverages and fruit juices	Total sugars intake in grams per day, averaged more than age 5–8 years
	Sugar-sweetened beverages intake in ounces per day (excluding 100% fruit juice), averaged more than age 5–8 years
	Fruit juices intake in ounces per day, averaged more than age 5–8 years using the procedure described in the manuscript to handle missing data
Caries variables	Indicator for any surfaces with cavitated caries or filling at the primary dentition (age 5) exam
	Indicator for any surfaces with non-cavitated caries or a filling at the primary dentition (age 5) exam
	Count of surfaces with non-cavitated caries at age 5 years
	Count of surfaces with cavitated caries or filling at age 5 years
	Count of surfaces with new non-cavitated or cavitated caries or filling from age 5 to 9 years (surfaces that transition from missing or sound to non-cavitated caries, cavitated caries, or fillings)
	Count of surfaces with new cavitated caries or filling from age 5 to 9 years (surfaces that transition from missing or sound to cavitated caries or fillings)
	Indicator for any surfaces with new non-cavitated or cavitated caries or filling from age 5 to 9 years (surfaces that transition from missing or sound to non-cavitated caries, cavitated caries, or fillings)
	Indicator for any surfaces with new cavitated caries or filling from age 5 to 9 years (surfaces that transition from missing or sound to cavitated caries or fillings)
Other confounding variables	Socio-economic status (SES) category at baseline, divided into three categories based on the table above
	Participant sex (1 = female, 0 = male)
	Age (in years) at mixed dentition exam (target age 9 years exam)
	Age interval (in years) between primary and mixed dentition exam (target ages 5 and 9 years)
	Mean daily tooth-brushing frequency, averaged more than age 5–8.5 years
	Composite water fluoride concentration (ppm), averaged more than age 5–8.5 years

2.2. Results

The original data set includes 198 subjects; the set under analysis includes 192 subjects [six subjects dropped out due to socio-economic status (SES) missing data]. Additional variables have been created:

- Count of surfaces with non-cavitated or cavitated caries or filling at age 5 years as sum of the two original variables.
- Any surfaces with non-cavitated or cavitated caries or filling at age 5 years, as an indicator from the related count variable.
- Count of surfaces with new non-cavitated caries from age 5 to 9 years, as difference from the two original variables.
- Any surfaces with new non-cavitated caries from age 5 to 9 years, as an indicator from the related count variable.
- Any surfaces with no caries, only non-cavitated caries, only cavitated caries, both types (including filling) at age 5 years, as combinations of the related indicators.
- Children with no caries (any type) at age 5 years.

Table 1: Main exposures and potential confounders

Variable	n	Mean	SD	Min	Max

Total sugar intake (g/day) (average)	192	114.51	27.31	53.23	215.98
Sugar-sweetened beverages intake (oz/day) (average)	192	9.18	5.9	0	36.5
Sugar-sweetened beverages intake (mL/day) (average)	192	271.63	174.55	0	1079.45
Fruit juices intake (oz/day) (average)	192	2.93	2.67	0	17.75
Fruit juices intake (mL/day) (average)	192	86.63	79.04	0	524.94
Participant sex (1 = female, 0 = male)	192	0.55	0.5	0	1
Age (in years) at mixed dentition exam (age 9 years)	192	9.18	0.72	7.79	11.34
Age interval between primary and mixed dentition exam (5 and 9 years)	192	4.05	0.78	2.18	6.07
Mean daily tooth-brushing frequency (average)	192	1.51	0.49	0.5	3
Composite water fluoride concentration (ppm) (average)	192	0.83	0.36	0.11	2.24

Table 2: Caries outcomes as counts

Variable	n	Mean	SD	Min	Max
Count of surfaces with non-cavitated caries (age 5 years)	192	0.49	1.58	0	15
Count of surfaces with cavitated caries or filling (age 5 years)	192	1.04	2.9	0	18
Count of surfaces with non-cavitated or cavitated caries or filling (age 5 years)	192	1.53	3.7	0	24
Count of surfaces with new non-cavitated caries (age 5–9 years)	192	0.69	1.34	0	7
Count of surfaces with new cavitated caries or filling (age 5–9 years)	192	0.94	1.76	0	13
Count of surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)	192	1.63	2.35	0	13

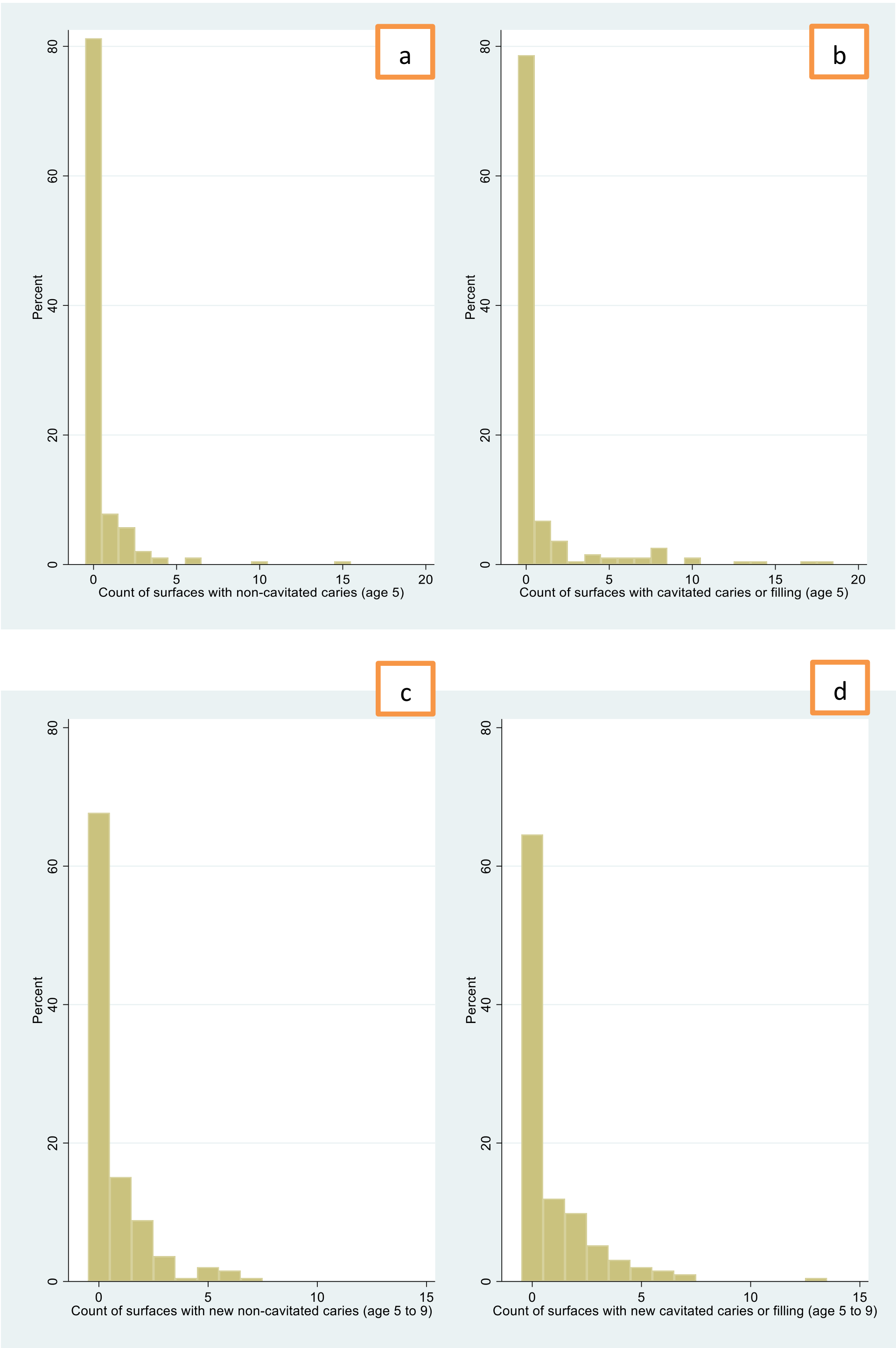
Table 3: Caries outcomes as indicators (the mean represents the proportion of counts > 0)

Variable	n	Mean	SD	Min	Max
Any surfaces with non-cavitated caries (age 5 years)	192	0.19	0.39	0	1
Any surfaces with cavitated caries or a filling (age 5 years)	192	0.21	0.41	0	1
Any surfaces with non-cavitated or cavitated caries or filling (age 5)	192	0.3	0.46	0	1
Any surfaces with new non-cavitated caries (age 5–9 years)	192	0.32	0.47	0	1
Any surfaces with new cavitated caries or filling (age 5–9 years)	192	0.35	0.48	0	1
Any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)	192	0.51	0.5	0	1

Table 4: Any surfaces at age 5 years by type of caries combined

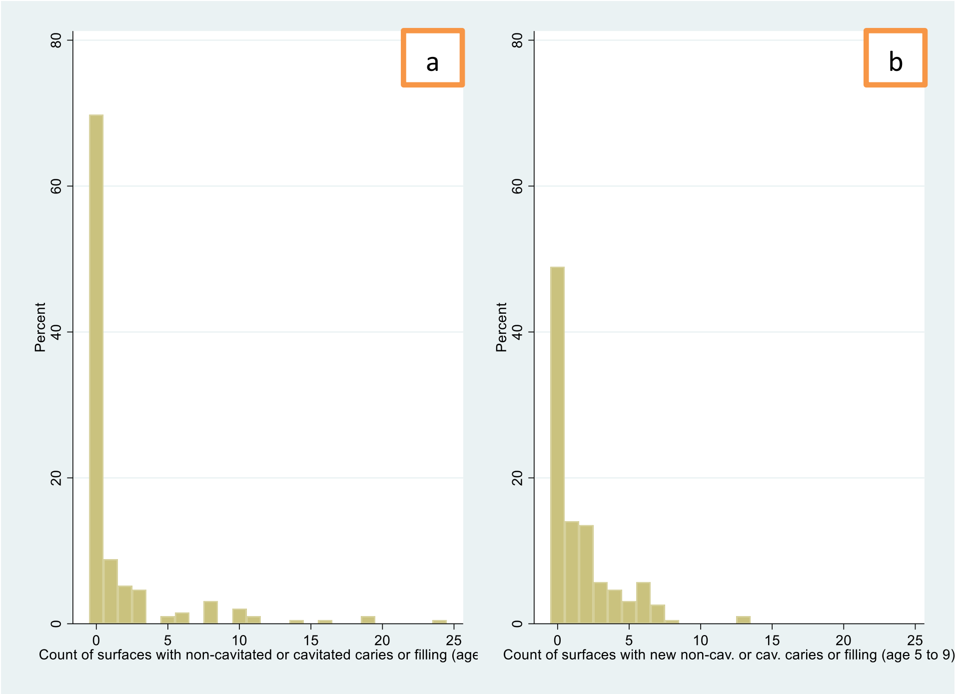
Levels	Frequency	Percentage	Cumulative
No caries	134	69.79	69.79
Only non-cavitated (NC) caries	17	8.85	78.65

Only cavitated (C) caries	22	11.46	90.1
Both NC and C caries	19	9.9	100



(a) Non-cavitated caries at 5 years; (b) cavitated caries or fillings at 5 years; (c) non-cavitated caries increment at 9 years; (d) cavitated caries or fillings increment at 9 years.

Figure 1: Frequency distribution of caries outcomes as counts



(a) Non-cavitated or cavitated caries or fillings at age 5 years; (b) new non-cavitated or cavitated caries or fillings increment from age 5 to 9 years.

Figure 2: Frequency distribution of caries outcomes as counts

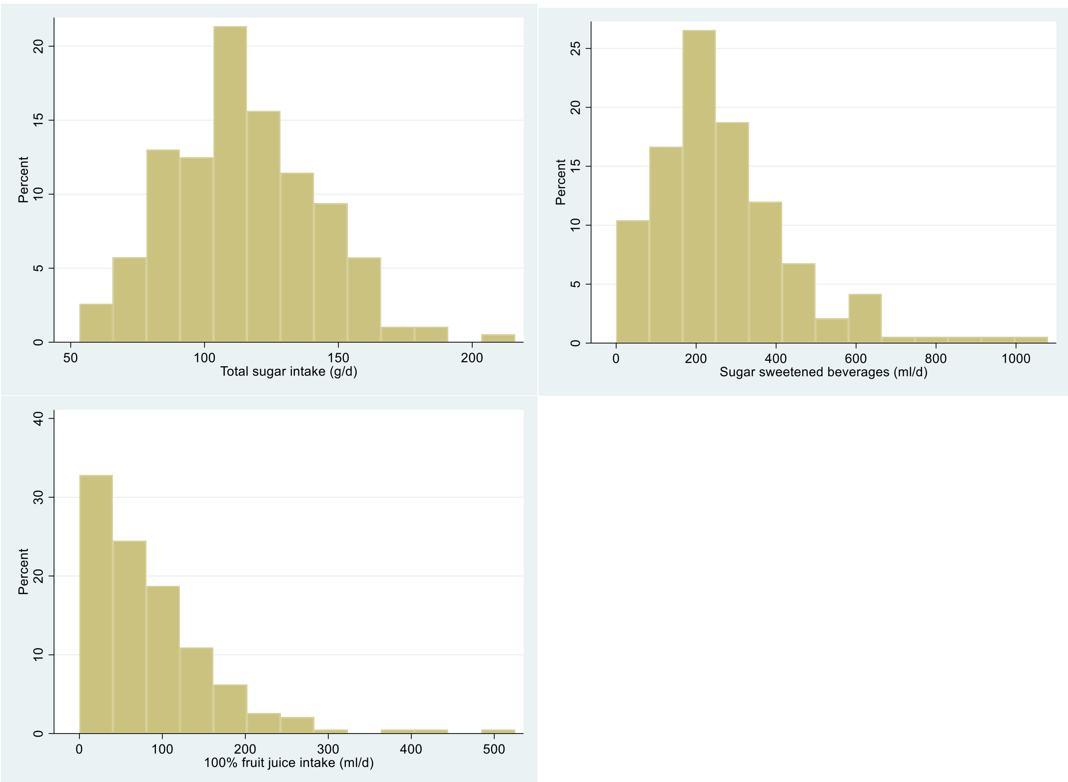


Figure 3: Observed distributions of total sugars intake (g/day), sugar-sweetened beverages intake (mL/day) and fruit juices intake (mL/day)

2.2.1. Association between total sugars intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (indicator)

The association of total sugars intake with the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' was tested first in univariable (Figure 4) and then in multivariable models (Figure 5), including the following potential confounders (Figure 4): gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level.

The final model was identified by comparing goodness-of-fit statistics [pseudo R-squared, Akaike's information criteria (AIC), Bayesian information criteria (BIC); Table 6].

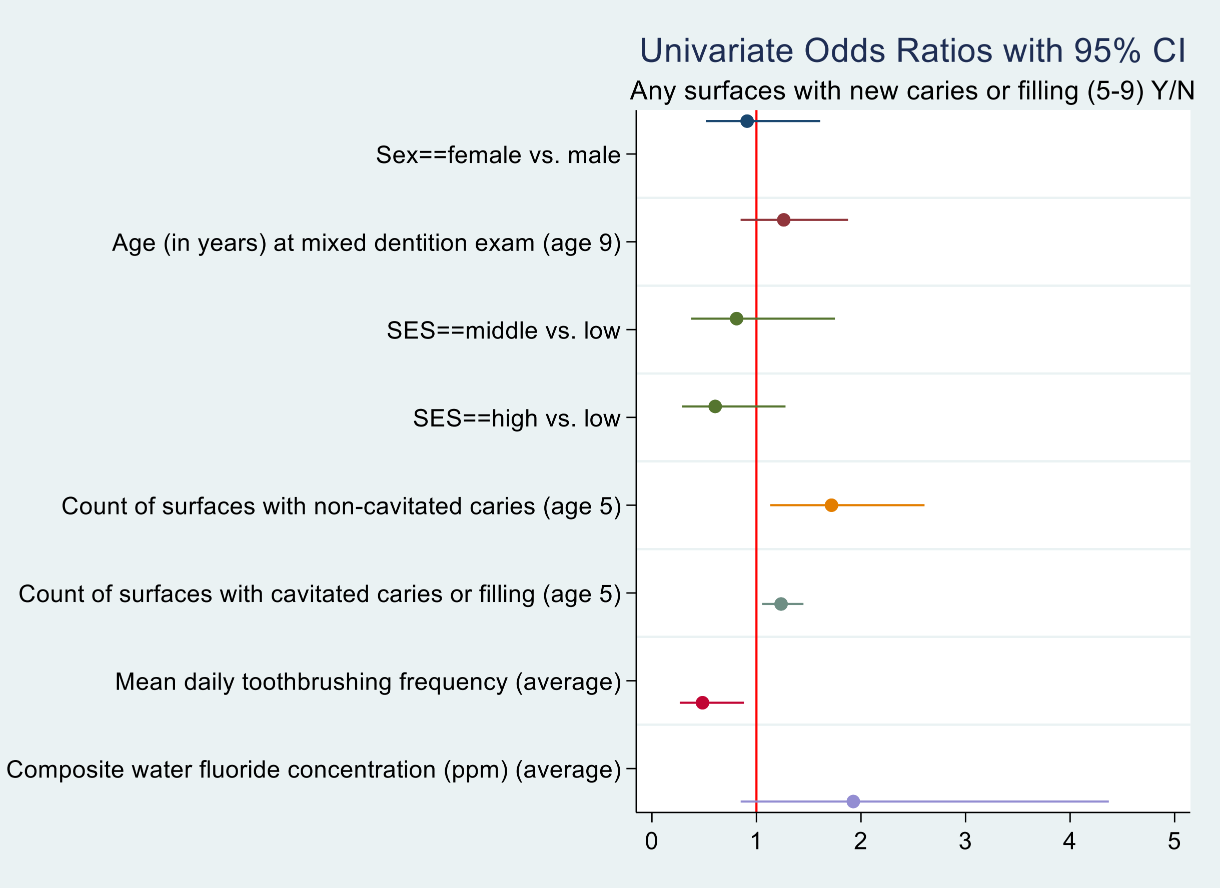
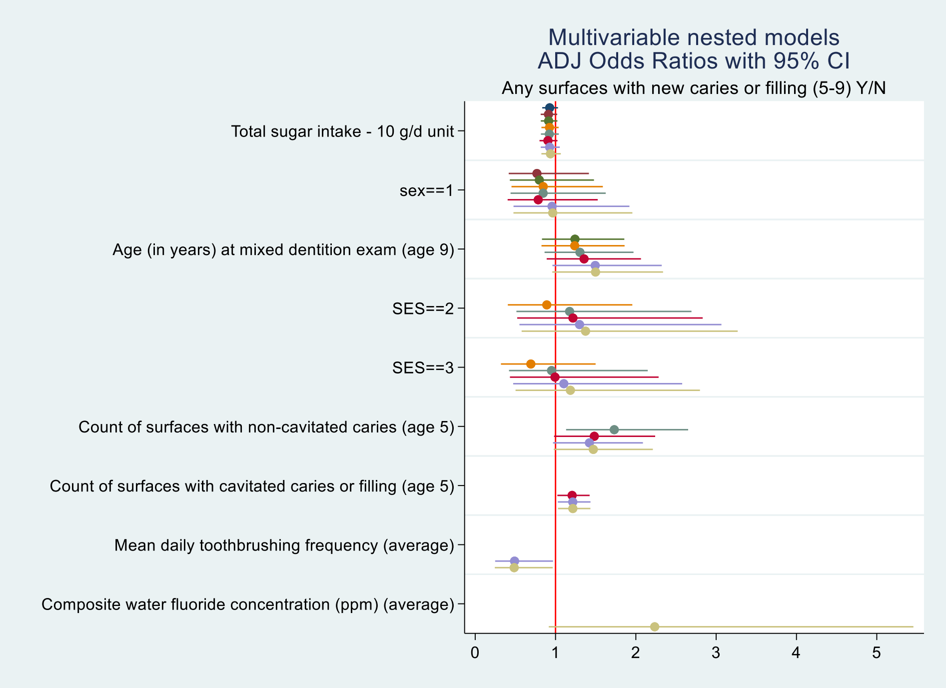


Figure 4: Univariate associations of outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' with potential and confounders – crude univariate ORs with 95% CIs



Odds ratios are mutually adjusted and reported with their 95% CIs. For each covariate it is possible to appreciate how further adjustments impact the point estimate and its precision.

Figure 5: Modelling strategy – the colour coding represents subsequent nested models, where covariates were added one at a time after the main exposure

Table 5: ORs with 95% CIs are reported for each model specification (ITS models across columns) and each covariate (along rows); the effect estimates of **total sugars intake** with increasing level of adjustment are highlighted in light blue; covariates associated with the outcome are highlighted in yellow; goodness-of-fit statistics are reported at the bottom of the table and compared for model selection (green values)

Any surfaces with new caries or filling (5-9) (Y/N)	ITS_0	ITS_C	ITS_1	ITS_2	ITS_3	ITS_4	ITS_5	ITS_6	ITS_A	ITS_F
	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95
Total sugar intake - 10 g/d unit		0.93	0.91	0.91	0.93	0.92	0.91	0.93	0.94	0.93
		0.83,1.03	0.82,1.02	0.82,1.02	0.82,1.04	0.82,1.04	0.80,1.02	0.82,1.05	0.82,1.07	0.83,1.04
sex==female			0.77	0.8	0.85	0.85	0.78	0.96	0.97	
			0.42,1.42	0.43,1.48	0.45,1.59	0.44,1.62	0.40,1.53	0.48,1.92	0.48,1.96	
Age (in years) at mixed dentition exam (age 9)				1.24	1.24	1.3	1.36	1.49	1.5	1.47
				0.83,1.86	0.83,1.86	0.86,1.97	0.89,2.06	0.96,2.32	0.96,2.34	0.96,2.25
SES==2					0.89	1.18	1.22	1.3	1.37	
					0.41,1.96	0.51,2.69	0.52,2.83	0.55,3.07	0.58,3.27	
SES==3					0.69	0.95	0.99	1.1	1.19	
					0.32,1.50	0.42,2.15	0.43,2.28	0.47,2.58	0.50,2.80	
Count of surfaces with non-cavitated caries (age 5)						1.73	1.48	1.42	1.47	1.41
						1.13,2.65	0.98,2.24	0.97,2.09	0.98,2.21	0.96,2.07
Count of surfaces with cavitated caries or filling (age 5)							1.21	1.21	1.22	1.21
							1.02,1.42	1.03,1.44	1.03,1.43	1.03,1.43
Mean daily toothbrushing frequency (average)								0.49	0.49	0.49
								0.25,0.97	0.24,0.96	0.25,0.93
Composite water fluoride concentration (ppm) (average)									2.24	
									0.92,5.46	
_cons	1.04	2.47	3.45	0.45	0.47	0.2	0.16	0.12	0.05	0.15
	0.79,1.38	0.71,8.55	0.79,15.02	0.01,25.51	0.01,29.83	0.00,14.50	0.00,12.32	0.00,10.49	0.00,5.19	0.00,8.89
N	192	192	192	192	192	192	192	192	192	192
Pseudo R-squared	0%	1%	1%	1%	2%	6%	8%	10%	11%	10%
AIC	268	268	269	270	273	265	260	258	257	252
BIC	271	275	279	283	293	287	286	287	289	272

Table 6: Final model on the association between total sugars intake and the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=192

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
Total sugar intake - 10 g/d unit				
per unit	0.93	0.05	0.225	(0.83 to 1.04)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.47	0.32	0.076	(0.96 to 2.25)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.21	0.10	0.023	(1.03 to 1.43)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.41	0.28	0.079	(0.96 to 2.07)
Mean daily toothbrushing frequency (average)				
per unit	0.49	0.16	0.028	(0.25 to 0.93)

Table 7: Most adjusted model on the association between total sugars intake and the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' restricted to children with no caries at age 5 years. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=134

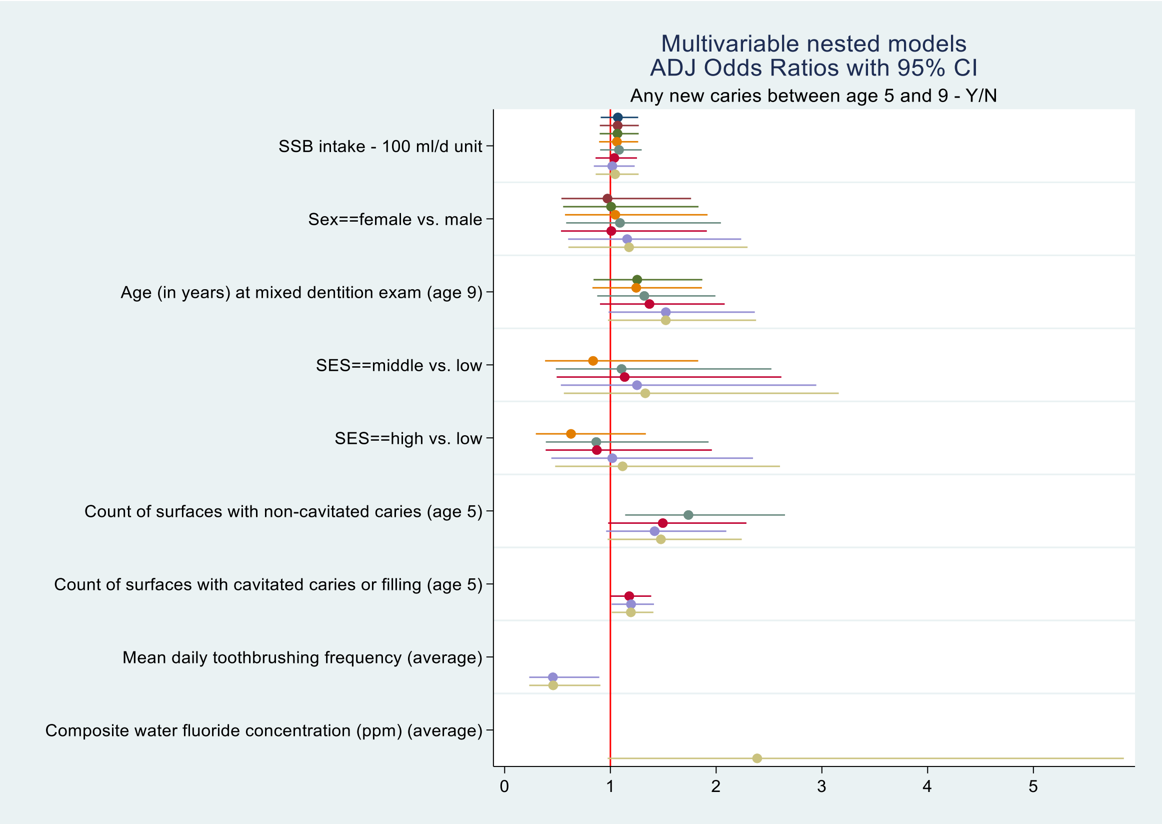
Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
Total sugar intake - 10 g/d unit				
per unit	0.88	0.07	0.109	(0.75 to 1.03)
Participant sex (1 = female, 0 = male)				
0*	1			
1	0.90	0.38	0.802	(0.39 to 2.06)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.39	0.36	0.202	(0.84 to 2.30)
Socioeconomic status category at baseline - 3 levels				
1.Low*	1			
2.Middle	2.04	1.15	0.204	(0.68 to 6.15)
3.High	2.34	1.29	0.122	(0.80 to 6.87)
Mean daily toothbrushing frequency (average)				
per unit	0.60	0.24	0.194	(0.28 to 1.30)
Composite water fluoride concentration (ppm) (average)				
per unit	2.58	1.44	0.090	(0.86 to 7.69)

* Baseline category

2.2.2. Association between sugar-sweetened beverages intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (indicator)

The association of sugar-sweetened beverages intake and the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' was tested first in univariable (Figure 4) and then in multivariable (Figure 6) models, including the following potential confounders (Figure 4): gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level.

The final model was identified by comparing goodness-of-fit statistics (pseudo R-squared, AIC, BIC; Table 9).



Odds ratios are mutually adjusted and reported with their 95% CIs. For each covariate it is possible to appreciate how further adjustments impact the point estimate and its precision.

Figure 6: Modelling strategy – the colour coding represents subsequent nested models, where covariates were added one at a time after the main exposure

Table 8: ORs with 95% CIs are reported for each model specification (ISSB models across columns) and each covariate (along rows); the effect estimates of sugar-sweetened beverages intake with increasing level of adjustment are highlighted in light blue; covariates associated with the outcome are highlighted in yellow; goodness-of-fit statistics are reported at the bottom of the table and compared for model selection (green values)

Any surfaces with new caries or filling (5-9) (Y/N)	ISSB_0	ISSB_C	ISSB_1	ISSB_2	ISSB_3	ISSB_4	ISSB_5	ISSB_6	ISSB_A	ISSB_F
	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95
SSB intake - 100 ml/d unit		1.07	1.07	1.07	1.06	1.08	1.04	1.02	1.04	1.01
		0.91,1.26	0.90,1.27	0.90,1.27	0.89,1.26	0.90,1.30	0.86,1.25	0.84,1.23	0.86,1.27	0.85,1.21
sex==female			0.97	1.01	1.05	1.09	1.01	1.16	1.18	
			0.54,1.76	0.55,1.83	0.57,1.92	0.58,2.05	0.53,1.91	0.60,2.24	0.60,2.30	
Age (in years) at mixed dentition exam (age 9)				1.25	1.24	1.32	1.37	1.52	1.52	1.48
				0.84,1.87	0.83,1.87	0.88,1.99	0.90,2.08	0.98,2.37	0.98,2.38	0.97,2.26
SES==2					0.84	1.11	1.14	1.25	1.33	
					0.38,1.83	0.48,2.52	0.49,2.62	0.53,2.95	0.56,3.16	
SES==3					0.63	0.87	0.87	1.02	1.12	
					0.30,1.34	0.39,1.93	0.39,1.96	0.44,2.35	0.48,2.60	
Count of surfaces with non-cavitated caries (age 5)						1.74	1.5	1.42	1.48	1.41
						1.14,2.65	0.98,2.29	0.96,2.10	0.97,2.24	0.96,2.08
Count of surfaces with cavitated caries or filling (age 5)							1.18	1.2	1.19	1.2
							1.00,1.39	1.01,1.41	1.01,1.41	1.01,1.41
Mean daily toothbrushing frequency (average)								0.46	0.46	0.47
								0.23,0.90	0.23,0.91	0.24,0.90
Composite water fluoride concentration (ppm) (average)									2.39	
									0.97,5.86	
_cons	1.04	0.87	0.88	0.11	0.15	0.05	0.04	0.04	0.02	0.07
	0.79,1.38	0.51,1.47	0.44,1.78	0.00,4.70	0.00,7.37	0.00,2.94	0.00,2.38	0.00,2.79	0.00,1.36	0.00,3.18
N	192	192	192	192	192	192	192	192	192	192
Pseudo R-squared	0%	0%	0%	1%	1%	5%	7%	9%	11%	9%
AIC	268	269	271	272	275	266	263	259	257	254
BIC	271	276	281	285	294	288	289	288	290	273

Table 9: Final model on the association between sugar-sweetened beverages intake and the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=192

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
SSB intake - 100 ml/d unit				
per unit	1.01	0.09	0.881	(0.85 to 1.21)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.48	0.32	0.072	(0.97 to 2.26)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.20	0.10	0.034	(1.01 to 1.41)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.41	0.28	0.080	(0.96 to 2.08)
Mean daily toothbrushing frequency (average)				
per unit	0.47	0.16	0.022	(0.24 to 0.90)

Table 10: Most adjusted model on the association between Sugar-sweetened beverages intake and the outcome 'any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' restricted to children with no caries at age 5 years. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=134

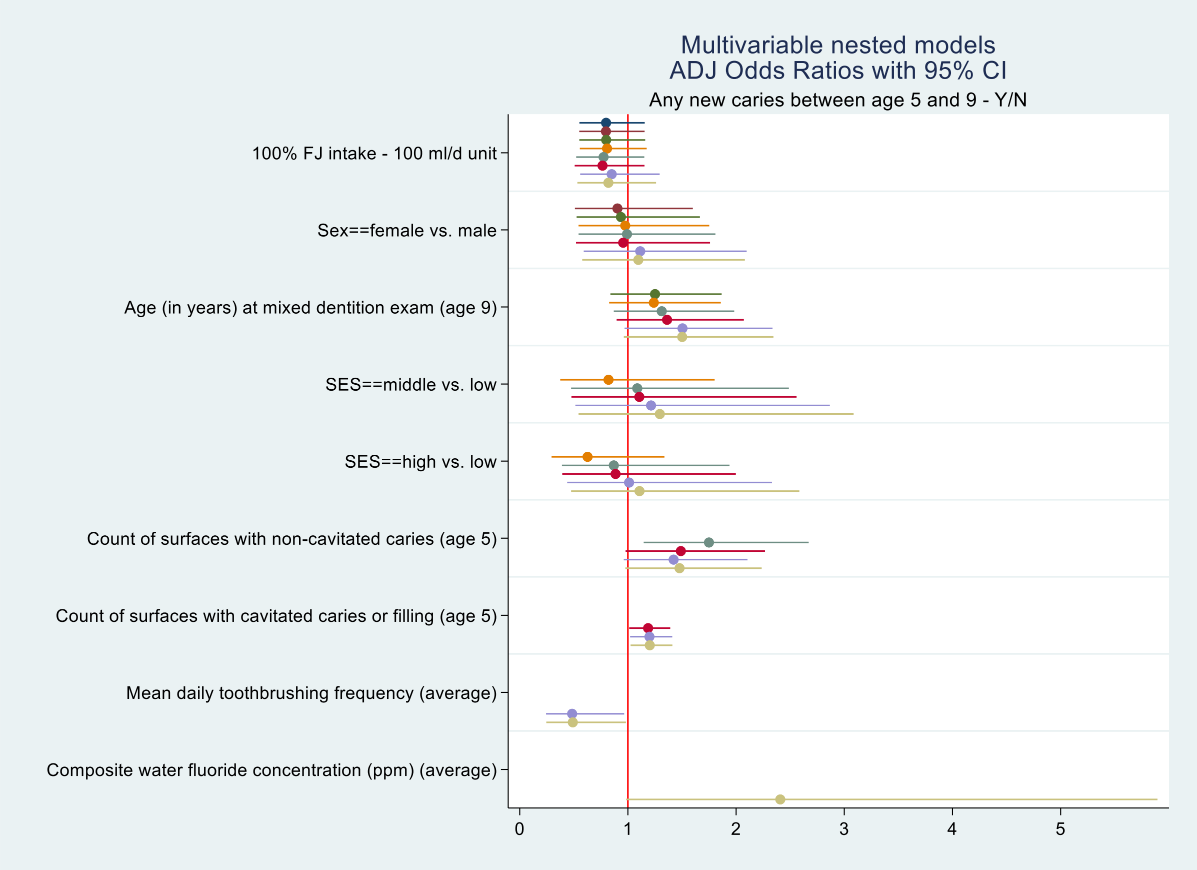
Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
SSB intake - 100 ml/d unit				
per unit	1.04	0.14	0.739	(0.81 to 1.35)
Participant sex (1 = female, 0 = male)				
0*	1			
1	1.24	0.49	0.579	(0.57 to 2.70)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.44	0.37	0.150	(0.88 to 2.38)
Socioeconomic status category at baseline - 3 levels				
1.Low*	1			
2.Middle	1.97	1.10	0.224	(0.66 to 5.88)
3.High	2.03	1.09	0.187	(0.71 to 5.81)
Mean daily toothbrushing frequency (average)				
per unit	0.54	0.21	0.116	(0.25 to 1.17)
Composite water fluoride concentration (ppm) (average)				
per unit	2.77	1.56	0.070	(0.92 to 8.34)

* Baseline category

2.2.3. Association between fruit juice intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (indicator)

The association between fruit juice intake and the outcome ‘any surfaces with new non-cavitated or cavitated caries or filling (age 5–9)’ was tested first in univariable and then in multivariable (Figure 7) models, including the following potential confounders (Figure 4): gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level.

The final model was identified by comparing goodness-of-fit statistics (pseudo R-squared, AIC, BIC; Table 12).



Odds ratios are mutually adjusted and reported with their 95% CIs. For each covariate it is possible to appreciate how further adjustments impact the point estimate and its precision.

Figure 7: Modelling strategy – the colour coding represents subsequent nested models, where covariates were added one at a time after the main exposure

Table 11: ORs with 95% CIs are reported for each model specification (I100FJ models across columns) and each covariate (along rows); the effect estimates of fruit juices (FJs) intake with increasing level of adjustment are highlighted in light blue; covariates associated with the outcome are highlighted in yellow; goodness-of-fit statistics are reported at the bottom of the table and compared for model selection (green values)

Any surfaces with new caries or filling (5-9) (Y/N)	I100FJ_0	I100FJ_C	I100FJ_1	I100FJ_2	I100FJ_3	I100FJ_4	I100FJ_5	I100FJ_6	I100FJ_A	I100FJ_F
	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95
100% FJ intake - 100 ml/d unit		0.8	0.8	0.8	0.81	0.78	0.77	0.85	0.82	0.83
		0.55,1.16	0.55,1.15	0.55,1.16	0.56,1.17	0.52,1.15	0.51,1.15	0.56,1.29	0.53,1.26	0.55,1.26
sex==female vs. male			0.9	0.94	0.98	0.99	0.96	1.11	1.1	
			0.51,1.60	0.53,1.67	0.54,1.75	0.54,1.81	0.52,1.76	0.59,2.10	0.58,2.08	
Age (in years) at mixed dentition exam (age 9)				1.25	1.24	1.31	1.36	1.5	1.5	1.47
				0.84,1.87	0.83,1.86	0.87,1.98	0.90,2.07	0.97,2.34	0.96,2.35	0.96,2.24
SES==middle vs. low					0.82	1.09	1.11	1.21	1.3	
					0.37,1.80	0.47,2.49	0.48,2.56	0.51,2.87	0.54,3.09	
SES==high vs. low					0.63	0.87	0.89	1.01	1.11	
					0.29,1.34	0.39,1.94	0.39,2.00	0.44,2.33	0.47,2.58	
Count of surfaces with non-cavitated caries (age 5)						1.75	1.49	1.42	1.48	1.42
						1.15,2.67	0.98,2.27	0.96,2.11	0.98,2.24	0.96,2.10
Count of surfaces with cavitated caries or filling (age 5)							1.19	1.2	1.2	1.2
							1.01,1.39	1.02,1.41	1.02,1.41	1.02,1.41
Mean daily toothbrushing frequency (average)								0.48	0.49	0.5
								0.24,0.97	0.25,0.98	0.26,0.96
Composite water fluoride concentration (ppm) (average)									2.41	
									0.98,5.90	
_cons	1.04	1.27	1.34	0.17	0.23	0.09	0.06	0.06	0.03	0.08
	0.79,1.38	0.83,1.94	0.78,2.29	0.00,7.14	0.00,11.39	0.00,4.99	0.00,3.56	0.00,3.56	0.00,1.89	0.00,3.77
N	192	192	192	192	192	192	192	192	192	192
Pseudo R-squared	0%	1%	1%	1%	2%	6%	8%	10%	11%	9%
AIC	268	269	271	271	274	265	261	259	257	253
BIC	271	275	280	284	293	288	287	288	289	273

Table 12: Final model on the association between FJs intake and the outcome 'Any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=192

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
100% FJ intake - 100 ml/d unit				
per unit	0.83	0.18	0.390	(0.55 to 1.26)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.47	0.32	0.078	(0.96 to 2.24)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.20	0.10	0.027	(1.02 to 1.41)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.42	0.28	0.078	(0.96 to 2.10)
Mean daily toothbrushing frequency (average)				
per unit	0.50	0.17	0.037	(0.26 to 0.96)

Table 13: Most adjusted model on the association between FJs intake and the outcome 'Any surfaces with new non-cavitated or cavitated caries or filling (age 5–9 years)' restricted to children with no caries at age 5 years. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate

Outcome variable: outI (Any surfaces with new caries or filling (5-9) (Y/N)), n=134

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
100% FJ intake - 100 ml/d unit				
per unit	0.80	0.20	0.372	(0.50 to 1.30)
Participant sex (1 = female, 0 = male)				
0*	1			
1	1.16	0.45	0.690	(0.55 to 2.46)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.42	0.36	0.170	(0.86 to 2.34)
Socioeconomic status category at baseline - 3 levels				
1.Low*	1			
2.Middle	1.89	1.06	0.255	(0.63 to 5.65)
3.High	2.01	1.07	0.193	(0.70 to 5.73)
Mean daily toothbrushing frequency (average)				
per unit	0.58	0.23	0.178	(0.27 to 1.28)
Composite water fluoride concentration (ppm) (average)				
per unit	2.73	1.52	0.071	(0.92 to 8.13)

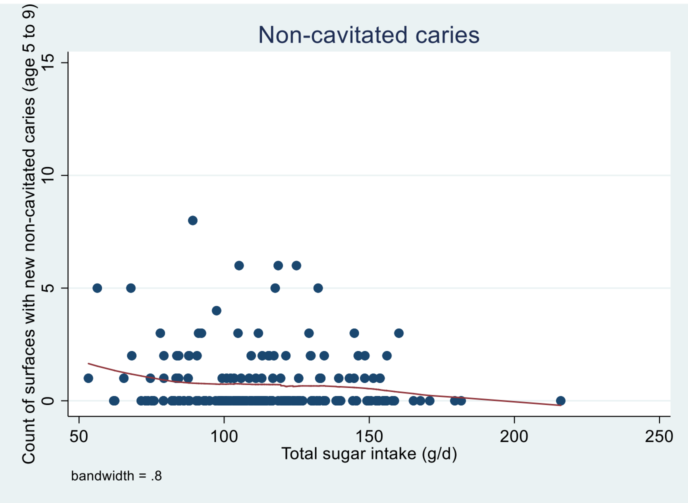
* Baseline category

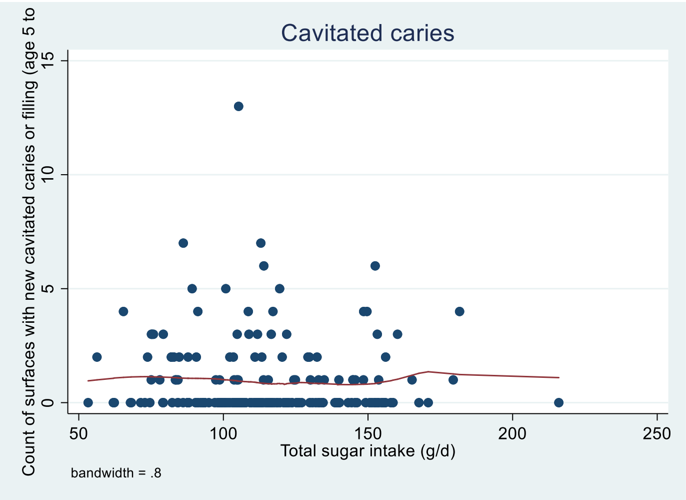
2.2.4. Association between total sugars intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (counts)

The association of total sugars intake with the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)' was tested first in univariable and then in multivariable models (applying the same modelling strategy as per indicators), including the following potential confounders: gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level. The same association was explored also considering only new non-cavitated caries as outcome.

Four different models (Poisson, negative binomial, zero-inflated Poisson, zero-inflated negative binomial) were tested to identify the approach that could best address the issues of overdispersion and excess zeros.

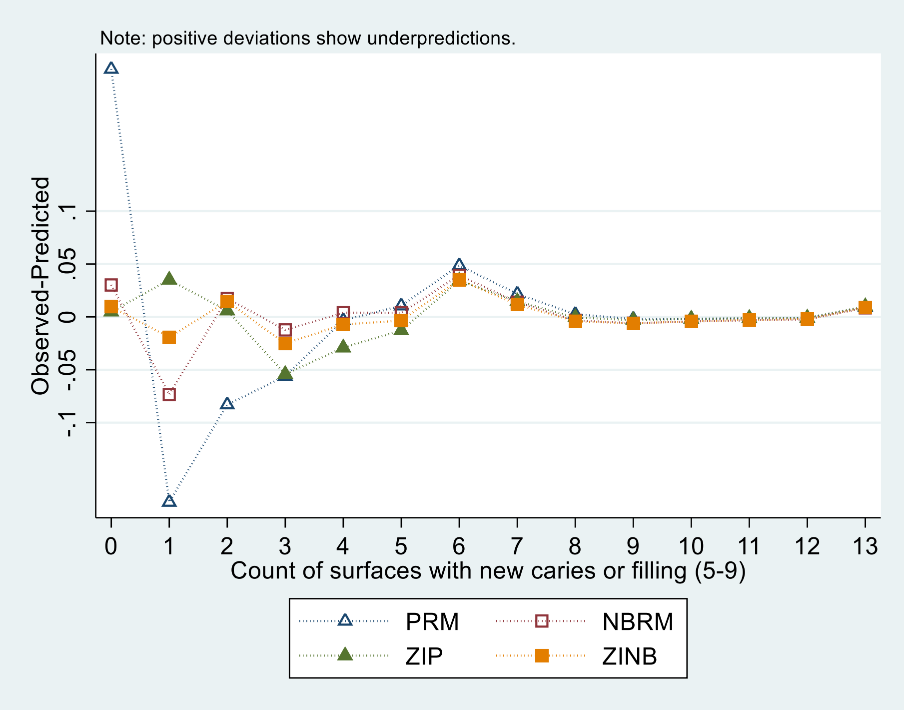
Also, different covariates were tested to account for the latent class of excess zeros (variable that could explain the excess, i.e. account for a fraction of subjects who did not developed caries at 9). The final model was identified by comparing goodness-of-fit (GoF) statistics (pseudo R-squared, AIC, BIC) and evaluating differences between observed and average estimated probabilities for each count (Figure 9).





The red line is a locally weighted regression line (smoothed) that captures the crude relationship between the dependent and independent variables without making assumptions on its shape (non-parametric).

Figure 8: Scatterplot of 'count of surfaces with new non-cavitated caries or fillings (age 5–9 years)' by total sugar intake and 'count of surfaces with new cavitated caries or fillings (age 5–9 years)' by total sugar intake



PRM = Poisson; NBRM = negative binomial; ZIP = zero-inflated Poisson; ZINB = zero-inflated negative binomial. The ZINB model shows the best fit (also from comparison of GoF statistics) as most differences are close to the zero line.

Figure 9: Differences between observed and average estimated probabilities for each count (1 to 13, observed outcome range)

Table 14: Final model on the association between total sugars intake and the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (Children with no caries at age 5 years) is not shown (p = 0.160)

Outcome variable: outC (Count of surfaces with new caries or filling (5-9)), n=192

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Total sugar intake - 10 g/d unit				
per unit	0.97	0.03	0.383	(0.91 to 1.04)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.34	0.20	0.051	(1.00 to 1.80)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.08	0.03	0.007	(1.02 to 1.14)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.06	0.04	0.130	(0.98 to 1.14)
Mean daily toothbrushing frequency (average)				
per unit	0.67	0.14	0.051	(0.45 to 1.00)

Table 15: Final model on the association between total sugars intake and the outcome 'count of surfaces with new non-cavitated caries only (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (children with no caries at age 5 years) is not shown (p = 0.120)

```
Outcome variable: d1_count_5_9 (Count of surfaces with new non-cavitated caries (age
>  5 to 9)), n=192
```

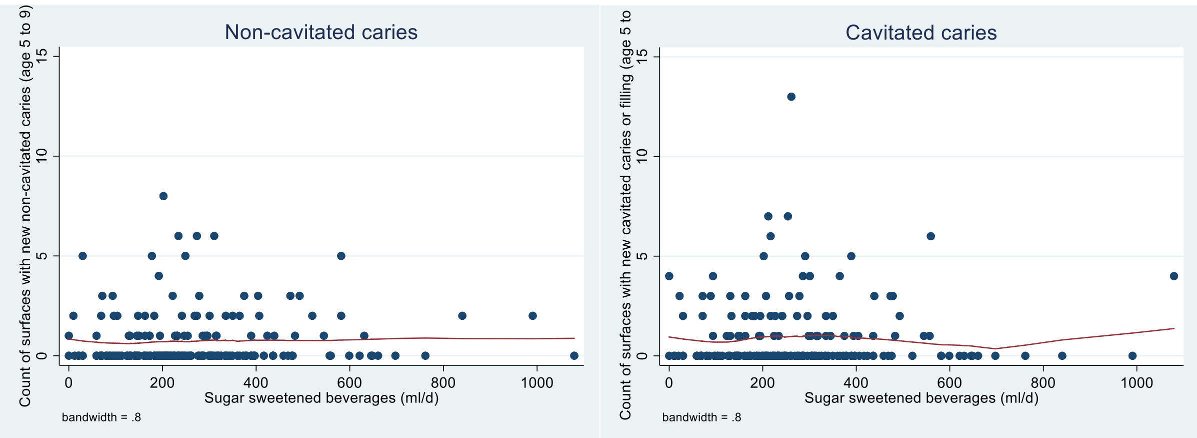
Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Total sugar intake - 10 g/d unit				
per unit	0.94	0.04	0.177	(0.85 to 1.03)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.34	0.26	0.133	(0.91 to 1.98)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.02	0.04	0.525	(0.95 to 1.10)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.09	0.05	0.063	(1.00 to 1.20)
Mean daily toothbrushing frequency (average)				
per unit	0.92	0.25	0.754	(0.53 to 1.58)

2.2.5. Association between sugar-sweetened beverages intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (counts)

The association of sugar-sweetened beverages intake with the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)' was tested first in univariable and then in multivariable models (applying the same modelling strategy as per indicators), including the following potential confounders: gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level. The same association was explored also considering only new non-cavitated caries as outcome.

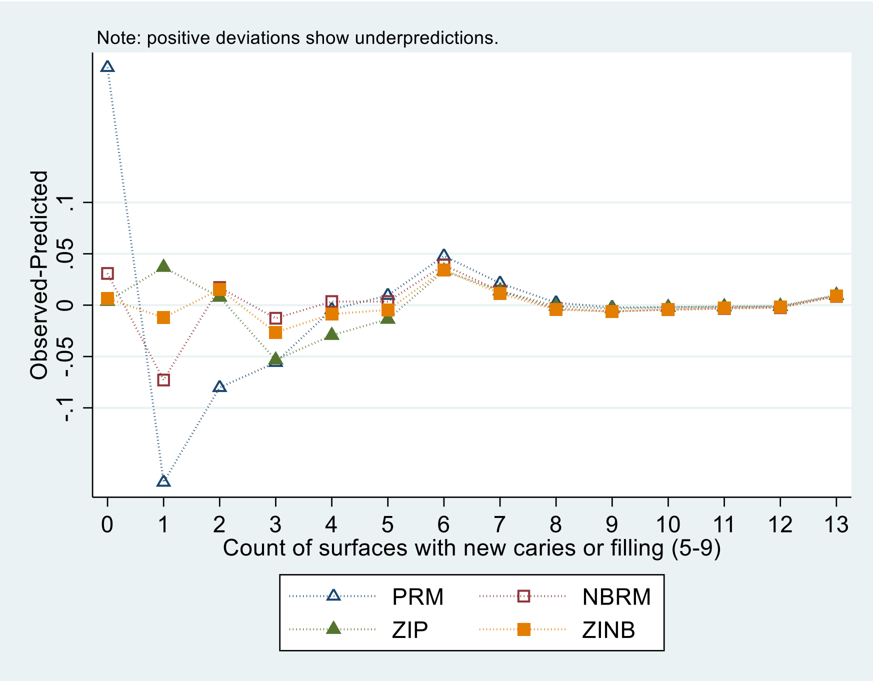
Four different models (Poisson, negative binomial, zero-inflated Poisson, zero-inflated negative binomial) were tested to identify the approach that could best address the issues of overdispersion and excess zeros.

Also, different covariates were tested to account for the latent class of excess zeros (variable that could explain the excess, i.e. account for a fraction of subjects who did not developed caries at 9). The final model was identified by comparing goodness-of-fit statistics (GoF) (pseudo R-squared, AIC, BIC) and evaluating differences between observed and average estimated probabilities for each count (Figure 11).



The red line is a locally weighted regression line (smoothed) that captures the crude relationship between the dependent and independent variables without making assumptions on its shape (non-parametric).

Figure 10: Scatterplot of 'count of surfaces with new non-cavitated (age 5–9 years)' by sugar-sweetened beverages intake and 'count of surfaces with new cavitated caries or fillings (age 5–9 years)' by total sugar intake



PRM = Poisson; NBRM = negative binomial; ZIP = zero-inflated Poisson; ZINB = zero-inflated negative binomial. The ZINB model shows the best fit (also from comparison of GoF statistics) as most differences are close to the zero line.

Figure 11: Differences between observed and average estimated probabilities for each count (1 to 13, observed outcome range)

Table 16: Final model on the association between sugar-sweetened beverages intake and the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (children with no caries at age 5 years) is not shown (p = 0.139)

Outcome variable: outC (Count of surfaces with new caries or filling (5-9)), n=192

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
SSB intake - 100 ml/d unit				
per unit	1.00	0.05	0.969	(0.90 to 1.10)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.33	0.20	0.060	(0.99 to 1.78)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.07	0.03	0.008	(1.02 to 1.13)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.06	0.04	0.110	(0.99 to 1.14)
Mean daily toothbrushing frequency (average)				
per unit	0.65	0.13	0.030	(0.44 to 0.96)

Table 17: Final model on the association between sugar-sweetened beverages intake and the outcome 'Count of surfaces with new non-cavitated caries only (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (children with no caries at age 5 years) is not shown (p = 0.104)

Outcome variable: d1_count_5_9 (Count of surfaces with new non-cavitated caries (age > 5 to 9)), n=192

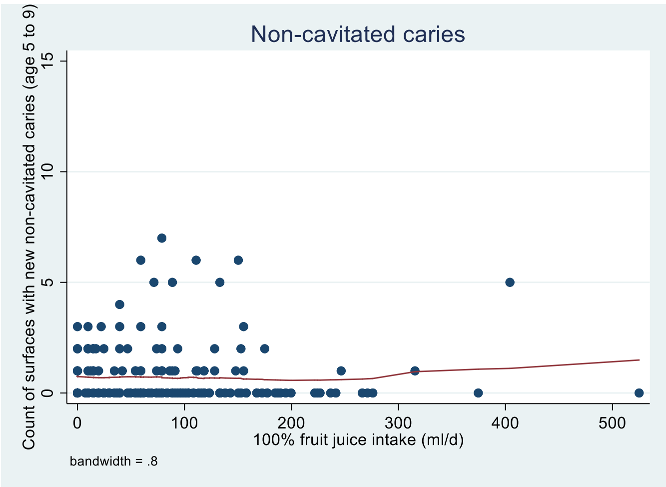
Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
SSB intake - 100 ml/d unit				
per unit	1.01	0.07	0.837	(0.88 to 1.17)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.36	0.27	0.123	(0.92 to 1.99)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.02	0.04	0.596	(0.95 to 1.09)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.10	0.05	0.048	(1.00 to 1.21)
Mean daily toothbrushing frequency (average)				
per unit	0.86	0.24	0.572	(0.50 to 1.47)

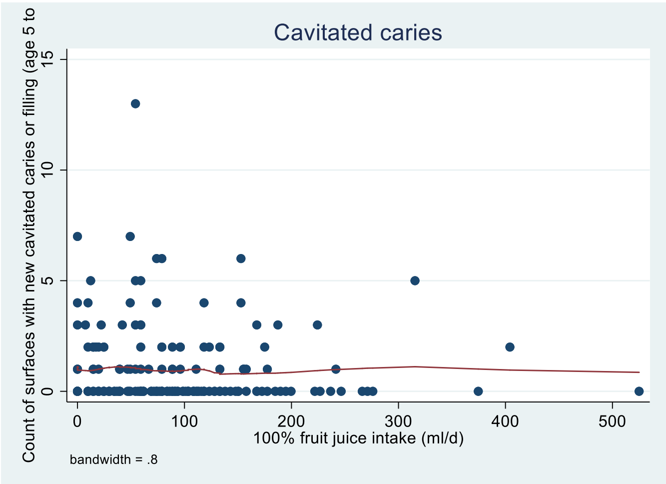
2.2.6. Association between fruit juices intake (average between age 5 and 8 years) and increment of dental caries between age 5 and 9 years (counts)

The association of FJs intake with the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)' was tested first in univariable and then in multivariable models (applying the same modelling strategy as per indicators), including the following potential confounders: gender, SES, age at mixed dentition exam, cavitated caries experience at 5 years old, non-cavitated caries experience at 5 years old; tooth-brushing frequency; and composite water fluoride level. The same association was explored also considering only new non-cavitated caries as outcome.

Four different models (Poisson, negative binomial, zero-inflated Poisson, zero-inflated negative binomial) were tested to identify the approach that could best address the issues of overdispersion and excess zeros.

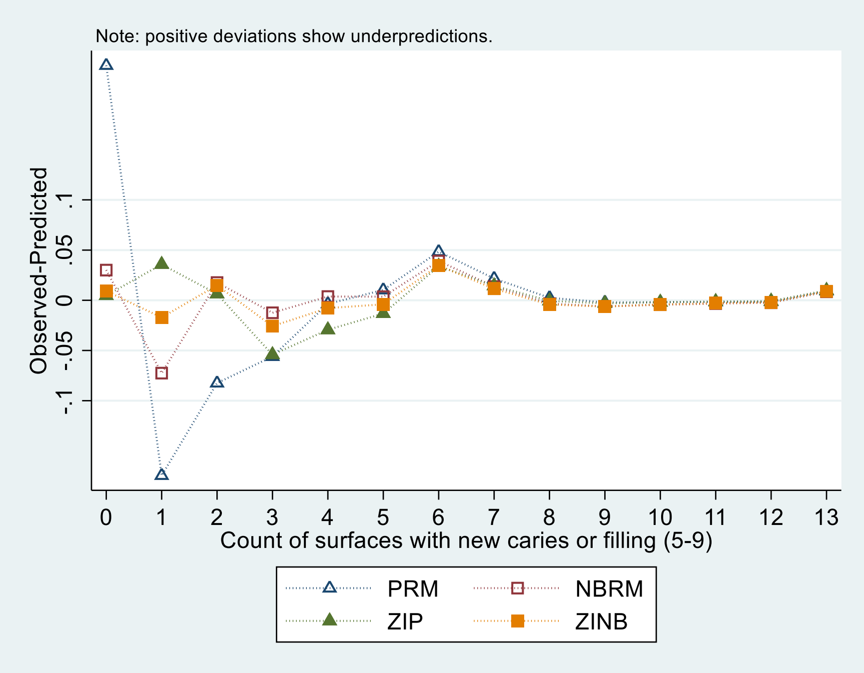
Also, different covariates were tested to account for the latent class of excess zeros (variable that could explain the excess, i.e. account for a fraction of subjects who did not developed caries at 9). The final model was identified by comparing goodness-of-fit statistics (GoF) (pseudo R-squared, AIC, BIC) and evaluating differences between observed and average estimated probabilities for each count (Figure 13).





The red line is a locally weighted regression line (smoothed) that captures the crude relationship between the dependent and independent variables without making assumptions on its shape (non-parametric).

Figure 12: Scatterplot of 'count of surfaces with new non-cavitated (age 5–9 years)' by fruit juices intake and 'count of surfaces with new cavitated caries or fillings (age 5–9 years)' by fruit juices intake



PRM = Poisson; NBRM = negative binomial; ZIP = zero-inflated Poisson; ZINB = zero-inflated negative binomial. The ZINB model shows the best fit (also from comparison of GoF statistics) as most differences are close to the zero line.

Figure 13: Differences between observed and average estimated probabilities for each count (1 to 13, observed outcome range)

Table 18: Final model on the association between FJs intake and the outcome 'count of surfaces with new non-cavitated or cavitated caries or fillings (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (children with no caries at age 5 years) is not shown (p = 0.331)

Outcome variable: outC (Count of surfaces with new caries or filling (5-9)), n=192

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
100% FJ intake - 100 ml/d unit				
per unit	0.96	0.12	0.779	(0.75 to 1.24)
Age (in years) at mixed dentition exam (age 9)				
per unit	1.33	0.20	0.057	(0.99 to 1.79)
Count of surfaces with cavitated caries or filling (age 5)				
per unit	1.08	0.03	0.008	(1.02 to 1.14)
Count of surfaces with non-cavitated caries (age 5)				
per unit	1.06	0.04	0.105	(0.99 to 1.15)
Mean daily toothbrushing frequency (average)				
per unit	0.66	0.14	0.043	(0.44 to 0.99)

Table 19: Final model on the association between fruit juices intake and the outcome 'count of surfaces with new non-cavitated caries only (age 5–9 years)'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (children with no caries at age 5 years) is not shown (p = 0.141)

```
Outcome variable: d1_count_5_9 (Count of surfaces with new non-cavitated caries (age 5 to 9))
> , n=192

Covariate                  exp(coef.)  Std. Err.     P>|z|     95% Conf. Interval

100% FJ intake - 100 ml/d unit
  per unit                       1.02        0.19     0.902        (0.71 to 1.48)
Age (in years) at mixed dentition exam (age 9)
  per unit                       1.39        0.29     0.110        (0.93 to 2.08)
Count of surfaces with cavitated caries or filling (age 5)
  per unit                       1.02        0.04     0.685        (0.94 to 1.10)
Count of surfaces with non-cavitated caries (age 5)
  per unit                       1.09        0.06     0.099        (0.98 to 1.21)
Mean daily toothbrushing frequency (average)
  per unit                       0.82        0.24     0.494        (0.47 to 1.44)
```

3. STRIP-2 study

3.1. Data set description

Variable Category	Description
	Subject identification number (148 subjects)
Sucrose intakes	**Sucrose intake in grams per day at age 3 years**
	Sucrose intake in grams per day at age 6 years
	Sucrose intake in grams per day at age 12 years
	Sucrose intake in grams per day at age 16 years
	Sucrose intake as E% at age 3 years
	Sucrose intake as E% at age 6 years
	Sucrose intake as E% at age 12 years
	Sucrose intake as E% at age 16 years
Caries variables	**d3mft primary – 3 years**
	d3mft primary – 6 years
	D3MFT permanent – 12 years
	D3MFT permanent – 16 years
	D3MFT permanent – 6 years
Other confounding variables	Participant sex (1 = female, 0 = male)
	Original intervention group (STRIP – prospective randomised trial)
	Caries-free age (years)
	Daily tooth-brushing frequency at age 3 years – Y/N
	Daily tooth-brushing frequency at age 6 years – Y/N
	Daily tooth-brushing frequency at age 12 years – Y/N
	Daily tooth-brushing frequency at age 16 years – Y/N

3.2. Results

The original data set includes 148 subjects; the set under analysis includes 128 subjects for the 3–6 years analyses (13 subjects dropped due to loss to follow-up + 3 negative increments + 2 missing intake data + 2 missing tooth-brushing data) and 81 subjects for the 12–16 years analyses (60 subjects dropped due to loss to follow-up + 7 missing intake data). Additional variables have been created:

- Count of new d3mft from age 3 to 6 years, as difference from the two original variables.
- Count of new D3MFT from age 12 to 16 years, as difference from the two original variables.
- Any new d3mft from age 3 to 6 years, as an indicator from the related count variable.
- Any new D3MFT from age 12 to 16 years, as an indicator from the related count variable.

Table 20: Main exposures and potential confounders

Variable	n	Mean	Std. Dev.	Min	Max
Sucrose intake – g/day at 3 years	144	28.52	11.3	7.4	65.9
Sucrose intake – g/day at 6 years	126	35.92	12.81	7.3	82
Sucrose intake – g/day at 12 years	91	34.68	14.73	7.1	78.8
Sucrose intake – g/day at 16 years	86	38.41	20.74	7.1	92.9
Participant sex (1 = female, 0 = male)	148	0.47	0.5	0	1
Caries-free age – years	148	2.67	1.47	0	5
Intervention group	148	1.46	0.5	1	2
Daily tooth-brushing Y/N – 3 years	148	0.02	0.14	0	1
Daily tooth-brushing Y/N – 6 years	133	0.21	0.41	0	1
Daily tooth-brushing Y/N – 12 years	114	0.42	0.5	0	1
Daily tooth-brushing Y/N – 16 years	88	0.61	0.49	0	1

Table 21: Caries outcomes as counts

Variable	n	Mean	SD	Min	Max
d3mft primary – 3 years	148	0.2	0.85	0	8
d3mft primary – 6 years	135	0.99	2.06	0	11
D3MFT permanent – 12 years	114	0.98	1.39	0	7
D3MFT permanent – 16 years	88	3.01	3.16	0	18
D3MFT permanent – 6 years	113	0.05	0.32	0	3
Counts of new d3mft between ages 3 and 6 years	132	0.82	1.89	0	11
Counts of new D3MFT between ages 12 and 16 years	88	2.14	2.47	0	14

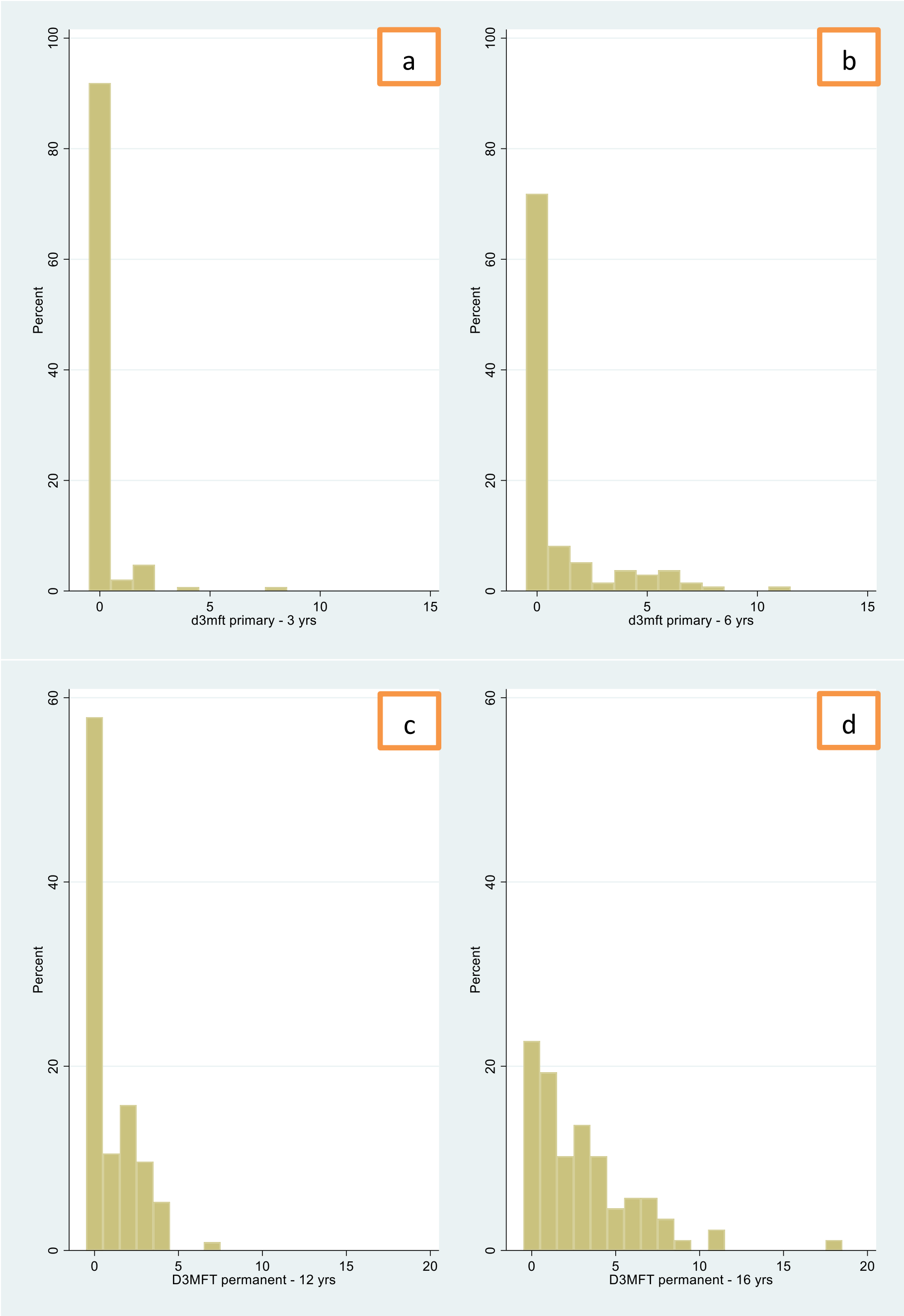
Table 22: Caries outcomes as indicators (the mean represents the proportion of counts > 0)

Variable	n	Mean	SD	Min	Max
Any new d3mft (age 3–6 years)	132	0.23	0.42	0	1
Any new D3MFT (age 12–16 years)	88	0.68	0.47	0	1

Table 23: Distributions of subjects' characteristics according to loss to follow-up at examination time (3, 6, 12 and 16 years)

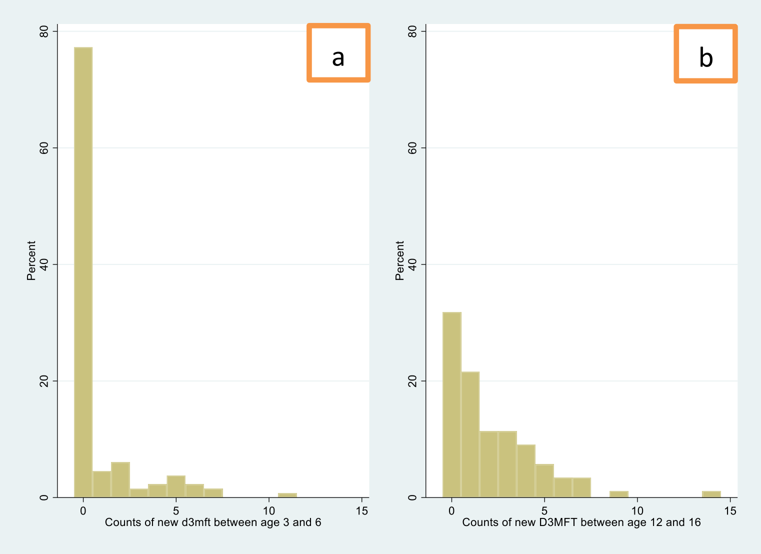
	Sucrose intake – g/day at 3 years	Sucrose intake – g/day at 6 years	Sucrose intake – g/day at 12 years	Sucrose intake – g/day at 16 years	Caries-free age – years	Sex (females %)	Daily tooth-brushing – 3 years	Daily tooth-brushing – 6 years	Daily tooth-brushing – 12 years	Daily tooth-brushing – 16 years
	d3mft primary dentition; 3 years (n = 148)									
N	144	126	91	86	148	148	148	133	114	88
Mean/ %	28.5	35.9	34.7	38.4	2.7	47%	2%	21%	42%	61%
SD	11.3	12.8	14.7	20.7	1.5	–	–	–	–	–
Min	7.4	7.3	7.1	7.1	0	–	–	–	–	–
Max	65.9	82	78.8	92.9	5	–	–	–	–	–

	d3mft primary dentition; 6 years (n = 135)									
N	133	124	91	86	135	135	135	133	113	87
Mean/ %	28.2	35.9	34.7	38.4	2.6	47%	2%	21%	42%	62%
SD	11.4	12.9	14.7	20.7	1.5	–	–	–	–	–
Min	7.4	7.3	7.1	7.1	0	–	–	–	–	–
Max	65.9	82	78.8	92.9	5	–	–	–	–	–
	D3MFT permanent dentition; 12 years (n = 114)									
N	114	109	84	81	114	114	114	111	114	88
Mean/ %	28.5	35.2	35.0	37.7	2.5	46%	3%	20%	42%	61%
SD	11.2	12.2	14.9	20.5	1.6	–	–	–	–	–
Min	7.4	16.3	7.1	7.1	0	–	–	–	–	–
Max	65.9	82	78.8	92.9	5	–	–	–	–	–
	D3MFT permanent dentition; 16 years (n = 88)									
N	88	85	81	81	88	88	88	87	88	88
Mean/ %	28.6	35.0	35.6	37.7	2.5	51%	1%	21%	44%	61%
SD	11.7	12.8	14.9	20.5	1.7	–	–	–	–	–
Min	7.4	16.3	7.1	7.1	0	–	–	–	–	–
Max	65.9	82	78.8	92.9	5	–	–	–	–	–



(a) d3mft primary – 3 years. (b) d3mft primary – 6 years. (c) D3MFT permanent – 12 years. (d) D3MFT permanent – 16 years.

Figure 14: Frequency distribution of caries outcomes as counts



(a) Counts of new d3mft between ages 3 and 6 years. (b) Counts of new D3MFT between ages 12 and 16 years.

Figure 15: Frequency distribution of caries outcomes as counts

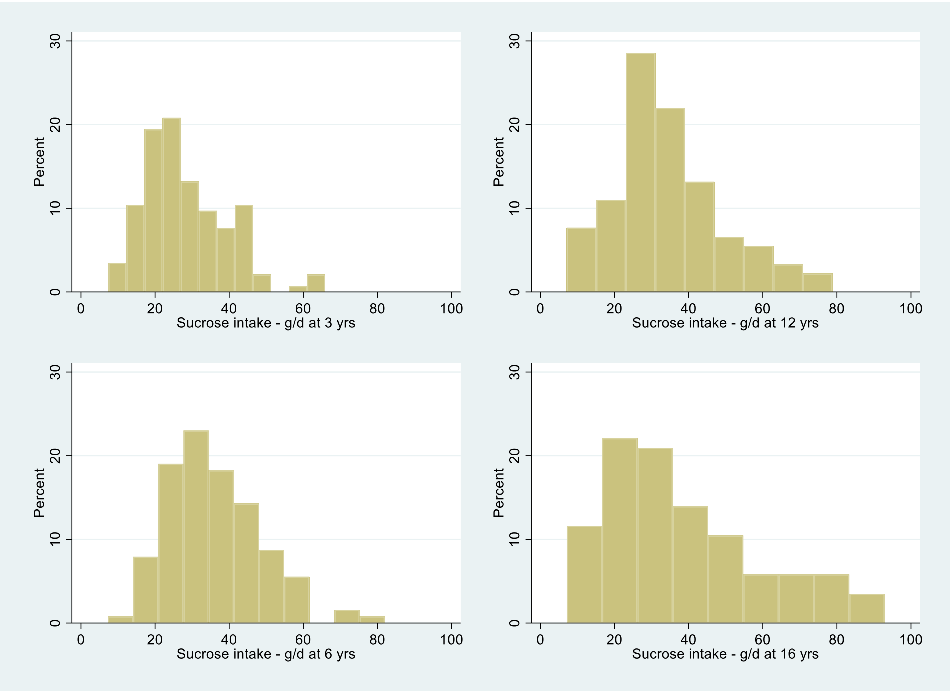


Figure 16: Frequency distribution of sucrose intake (g/day) at 3, 6, 12 and 16 years

3.2.1. Association between sucrose intake at age 3 years and increment of dental caries between ages 3 and 6 years (indicator)

The association of sucrose intake at age 3 years with the outcome 'any new d3mft (age 3–6 years)' was tested first in univariable (Figure 17) and then in multivariable (Figure 18) models, including the following potential confounders: intervention group, gender, caries experience at 3 years old, daily tooth-brushing (Y/N) at 3 years old. Daily tooth-brushing (Y/N) at 6 years old has also been tested, despite being collected cross-sectionally, as the indicator at age 3 years had a very low proportion.

The final model was identified by comparing GoF statistics (pseudo R-squared, AIC, BIC; Table 25).

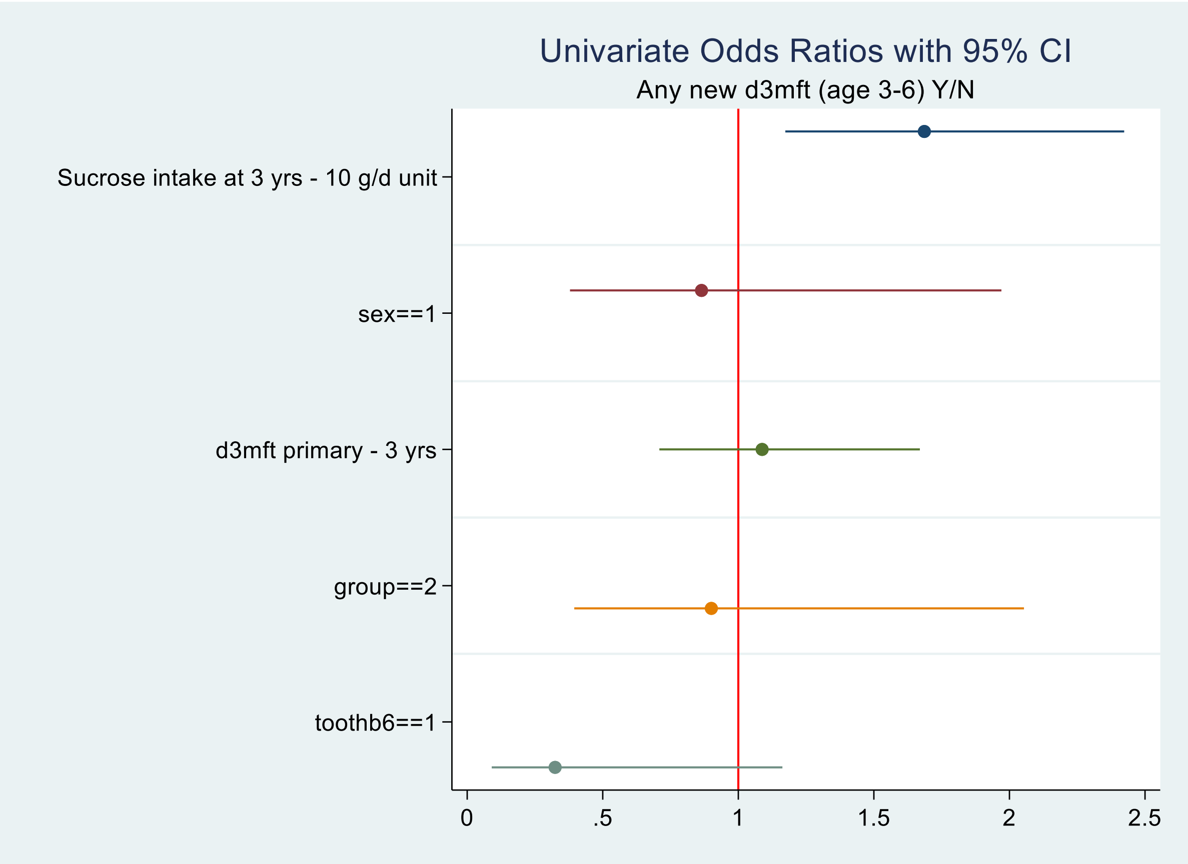
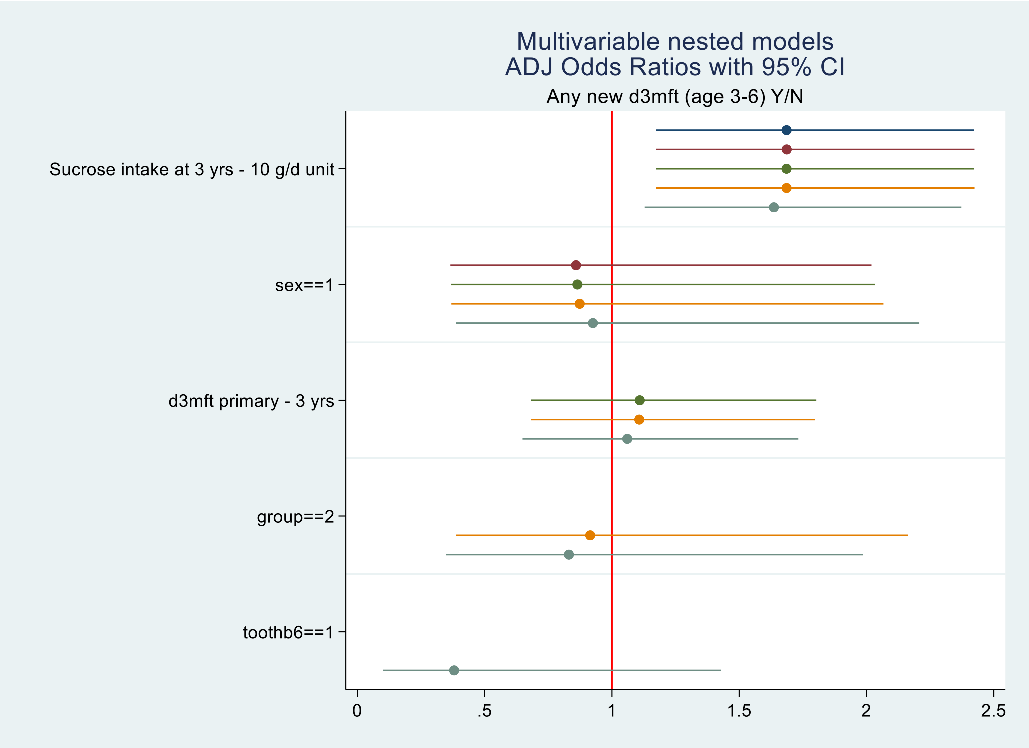


Figure 17: Univariate associations of outcome 'any new d3mft (age 3–6 years)' with main exposure and confounders – crude univariate ORs with 95% CIs



Odds ratios are mutually adjusted and reported with their 95% CIs. For each covariate it is possible to appreciate how further adjustments impact the point estimate and its precision.

Figure 18: Modelling strategy – the colour coding represents subsequent nested models, where covariates were added one at a time

Table 24: ORs with 95% CIs are reported for each model specification (I6 models across columns, I6_F: final model) and each covariate (along rows); the effect estimates of sucrose intake with increasing level of adjustment are highlighted in light blue; covariates associated with the outcome are highlighted in yellow; GoF statistics are reported at the bottom of the table and compared for model selection (green values)

Any new d3mft (age 3–6 years)	**I6_0**	**I6_C**	**I6_1**	**I6_2**	**I6_3**	**I6_A**	**I6_F**
	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95
Sucrose intake – g/day at 3 years		1.69	1.69	1.69	1.69	1.64	1.64
		1.17,2.42	1.17,2.42	1.17,2.42	1.17,2.42	1.13,2.37	1.13,2.37
sex==female			0.86	0.86	0.87	0.93	
			0.37,2.02	0.37,2.03	0.37,2.07	0.39,2.21	
d3mft primary – 3 years				1.11	1.11	1.06	
				0.68,1.80	0.68,1.80	0.65,1.73	
group==2					0.91	0.83	
					0.39,2.16	0.35,1.99	
toothb6==yes						0.38	0.38
						0.10,1.43	0.10,1.42

_cons	0.31	0.06	0.07	0.07	0.07	0.09	0.08
	0.20,0.46	0.02,0.22	0.02,0.24	0.02,0.24	0.02,0.26	0.02,0.36	0.02,0.29
N	128	128	128	128	128	128	128
Pseudo R-squared	0%	6%	6%	6%	6%	8%	8%
AIC	141	135	137	139	141	140	135
BIC	144	141	145	150	155	157	143

Table 25: Final model on the association between sucrose intake at age 3 years and the outcome 'any new d3mft (age 3–6 years)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. OR for sucrose intake is expressed per 10 g/day increase

Outcome variable: outI_3_6 (Any new d3mft (age 3-6)), n=128

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 3 yrs - 10 g/d unit				
per unit	1.64	0.31	0.009	(1.13 to 2.37)
Daily tooth-brushing Y/N - 6 yrs				
No*	1			
Yes	0.38	0.26	0.151	(0.10 to 1.42)

* Baseline category

Table 26: Non-linear model on the association between sucrose intake at age 3 years and the outcome 'any new d3mft (age 3–6 years)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. Categorical ORs are estimated per quartiles of sucrose intake (pseudo R-squared: 10%; AIC: 135; BIC: 149). A description of the exposure quartiles is reported after the model's results

Outcome variable: outI_3_6 (Any new d3mft (age 3-6)), n=128

Covariate	Odds Ratio	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 3 yrs - quartiles				
1*	1			
2	1.03	0.71	0.970	(0.26 to 4.01)
3	0.91	0.63	0.888	(0.23 to 3.54)
4	4.32	2.63	0.016	(1.31 to 14.25)
Daily tooth-brushing Y/N - 6 yrs				
No*	1			
Yes	0.38	0.26	0.153	(0.10 to 1.43)

* Baseline category

Sucrose intake at 3 yrs - quartiles	mean(sucrose3)	sd(sucrose3)	min(sucrose3)	max(sucrose3)
1	15.9	3.4	7.4	20.9
2	23.1	1.5	21.0	25.4
3	29.6	2.9	25.6	34.4
4	44.0	8.3	34.5	65.9

3.2.2. Association between sucrose intake at age 12 years and increment of dental caries between age 12 and 16 years (indicator)

The association of sucrose intake at age 12 years with the outcome 'any new D3MFT (age 12–16)' was tested first in univariable (Figure 19) and then in multivariable (Figure 20) models, including the following potential confounders: intervention group, gender, caries experience at 12 years of age, daily tooth-brushing (Y/N) at 12 years of age.

The final model was identified by comparing GoF statistics (pseudo R-squared, AIC, BIC; Table 28).

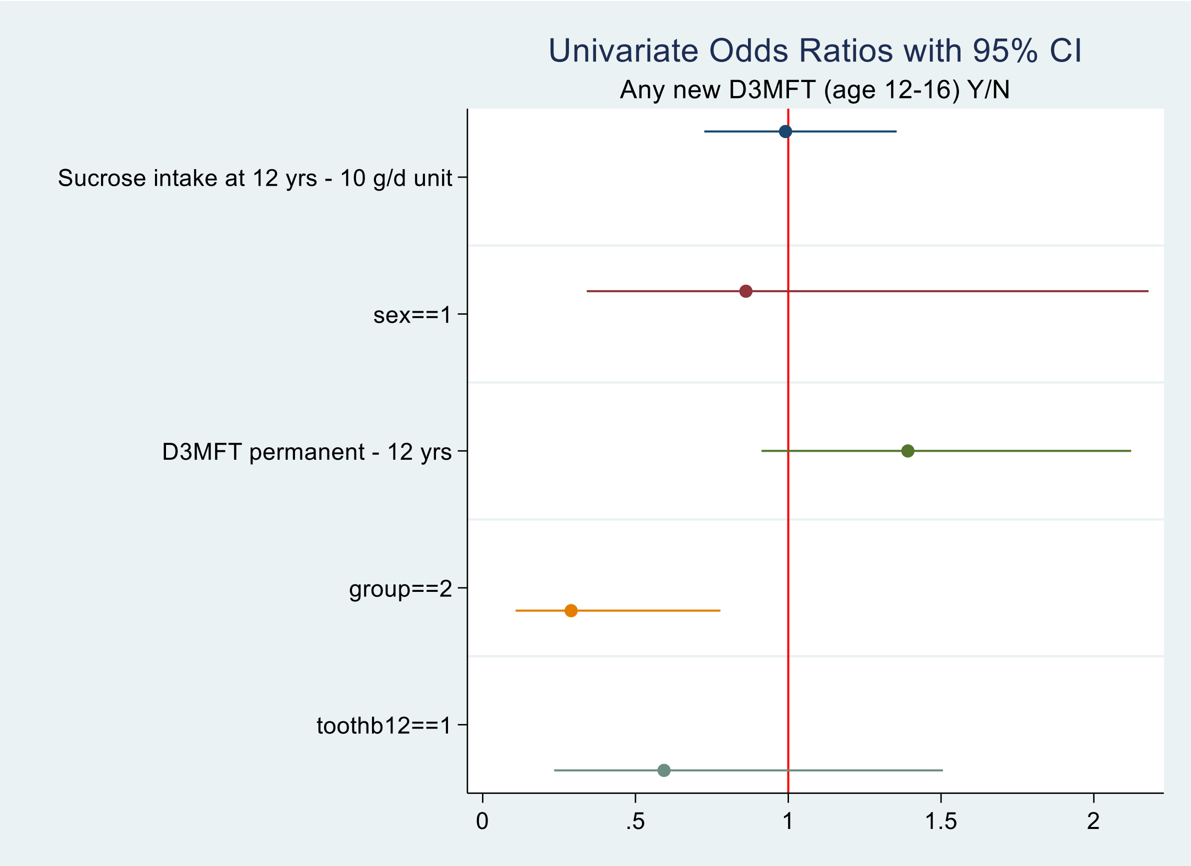
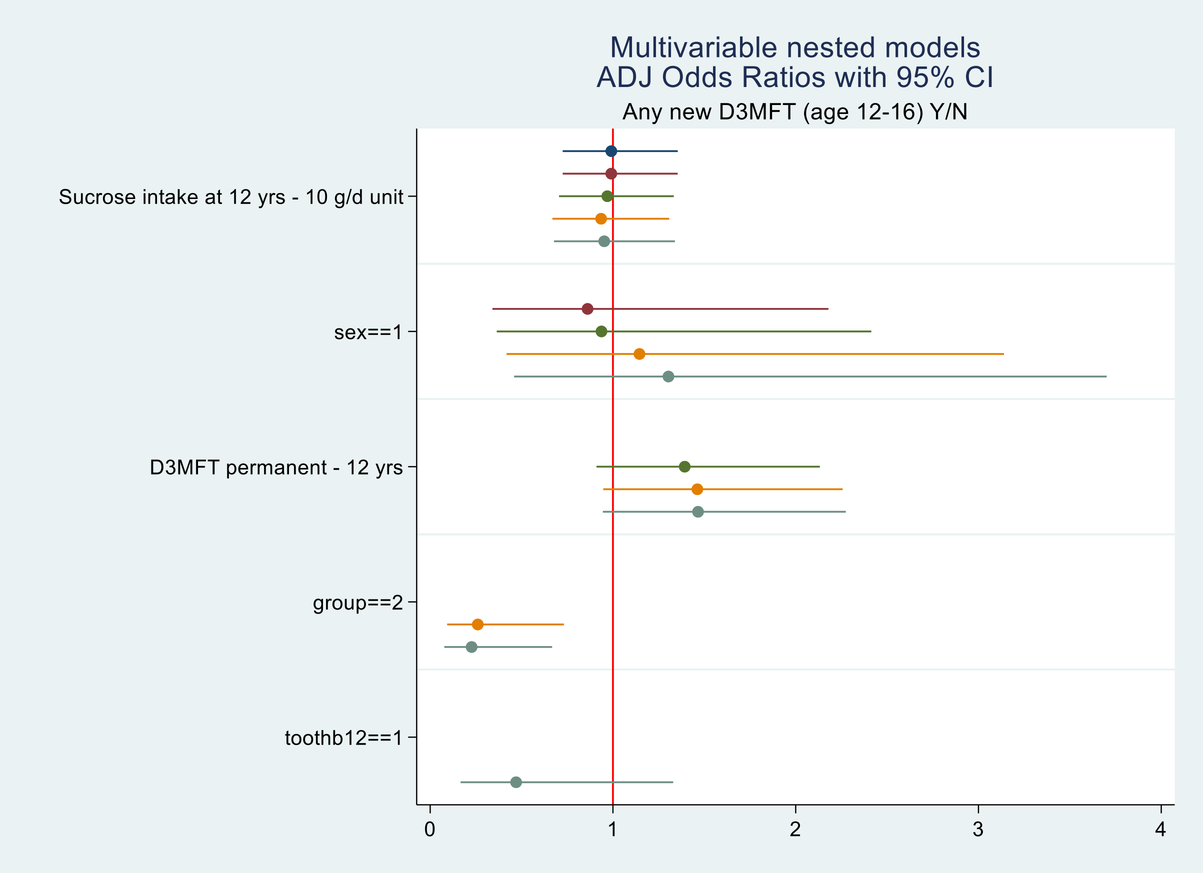


Figure 19: Univariate associations of outcome 'any new d3mft (age 12–16 years)' with main exposure and confounders – crude univariate ORs with 95% CIs



Odds ratios are mutually adjusted and reported with their 95% CIs. For each covariate it is possible to appreciate how further adjustments impact the point estimate and its precision.

Figure 20: Modelling strategy – the colour coding represents subsequent nested models, where covariates were added one at a time

Table 27: ORs with 95% CIs are reported for each model specification (I16 models across columns, I16_F final model) and each covariate (along rows); the effect estimates of **sucrose intake** with increasing level of adjustment are highlighted in light blue; covariates associated with the outcome are highlighted in yellow; GoF statistics are reported at the bottom of the table and compared for model selection (green values)

Any new D3MFT (age 12–16 years)	I16_0	I16_C	I16_1	I16_2	I16_3	I16_A	I16_F
	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95	b/ci95
Sucrose intake at 12 years – 10 g/day unit		0.99	0.99	0.97	0.94	0.95	0.95
		0.73,1.35	0.73,1.35	0.71,1.33	0.67,1.31	0.68,1.34	0.68,1.34
sex==female			0.86	0.94	1.15	1.3	
			0.34,2.18	0.36,2.41	0.42,3.14	0.46,3.70	
D3MFT permanent – 12 years				1.39	1.46	1.47	1.44
				0.91,2.13	0.95,2.26	0.94,2.27	0.93,2.23
group==2					0.26	0.23	0.24
					0.09,0.73	0.08,0.67	0.08,0.68

toothb12==yes						0.47	0.49
						0.17,1.33	0.18,1.36
_cons	2	2.06	2.24	1.77	3.67	5.03	5.5
	1.26,3.17	0.62,6.90	0.60,8.28	0.46,6.87	0.79,17.05	1.02,24.90	1.16,26.08
N	81	81	81	81	81	81	81
Pseudo R-squared	0%	0%	0%	3%	9%	11%	11%
AIC	105	107	109	108	103	103	102
BIC	108	112	116	118	115	118	113

Table 28: Final model on the association between sucrose intake at age 12 years and the outcome 'any new D3MFT (age 12–16)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The OR for sucrose intake is expressed per 10 g/day increase

```
Outcome variable: outI_12_16 (Any new D3MFT (age 12-16)), n=81
------------------------------------------------------------------------------
Covariate                Odds Ratio  Std. Err.     P>|z|     95% Conf. Interval
------------------------------------------------------------------------------
Sucrose intake at 12 yrs - 10 g/d unit
  per unit                     0.95       0.17     0.790          (0.68 to 1.34)
Intervention group
  1*                              1
  2                            0.24       0.13     0.007          (0.08 to 0.68)
D3MFT permanent - 12 yrs
  per unit                     1.44       0.32     0.100          (0.93 to 2.23)
Daily tooth-brushing Y/N - 12 yrs
  No*                             1
  Yes                          0.49       0.26     0.173          (0.18 to 1.36)
------------------------------------------------------------------------------
* Baseline category
```

Table 29: Non-linear model on the association between sucrose intake at age 12 years and the outcome 'any new D3MFT (age 12–16)'. Adjusted ORs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. Categorical ORs are estimated per quartiles of sucrose intake (pseudo R-squared: 15%; AIC: 102; BIC: 118). A description of the exposure quartiles is reported after the model's results

```
Outcome variable: outI_12_16 (Any new D3MFT (age 12-16)), n=81
------------------------------------------------------------------------------
Covariate                Odds Ratio  Std. Err.     P>|z|     95% Conf. Interval
------------------------------------------------------------------------------
Sucrose intake at 12 yrs - quartiles
  1*                              1
  2                            1.16       0.80     0.833          (0.30 to 4.50)
  3                            3.16       2.59     0.161         (0.63 to 15.75)
  4                            0.70       0.50     0.621          (0.17 to 2.84)
Intervention group
  1*                              1
  2                            0.23       0.13     0.009          (0.08 to 0.69)
D3MFT permanent - 12 yrs
  per unit                     1.38       0.31     0.145          (0.89 to 2.14)
Daily tooth-brushing Y/N - 12 yrs
  No*                             1
  Yes                          0.46       0.25     0.147          (0.16 to 1.31)
------------------------------------------------------------------------------
* Baseline category
```

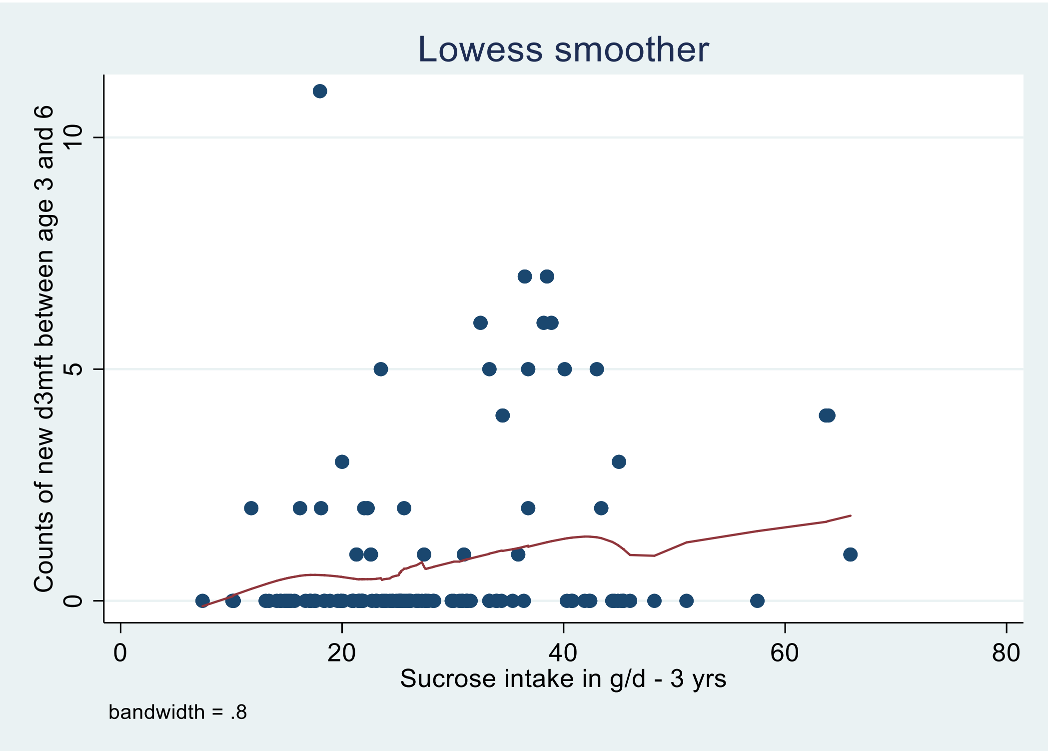
Sucrose intake at 12 yrs - quartiles	mean(sucro~12)	sd(sucro~12)	min(sucro~12)	max(sucro~12)
1	19.4	5.8	7.1	25.7
2	29.4	2.2	26.4	33.9
3	38.3	2.3	34.3	42.5
4	56.0	10.5	43.7	78.8

3.2.3. Association between sucrose intake at age 3 years and increment of dental caries between ages 3 and 6 years (counts)

The association of sucrose intake at age 3 years with the outcome 'counts of new d3mft between ages 3 and 6 years' was tested first in univariable and then in multivariable models (applying the same modelling strategy as per indicators), including the following potential confounders: intervention group, gender, caries experience at 3 years old, daily tooth-brushing (Y/N) at 3 years old. Daily tooth-brushing (Y/N) at 6 years old has also been tested, despite being collected cross-sectionally, as the indicator at age 3 had a very low proportion.

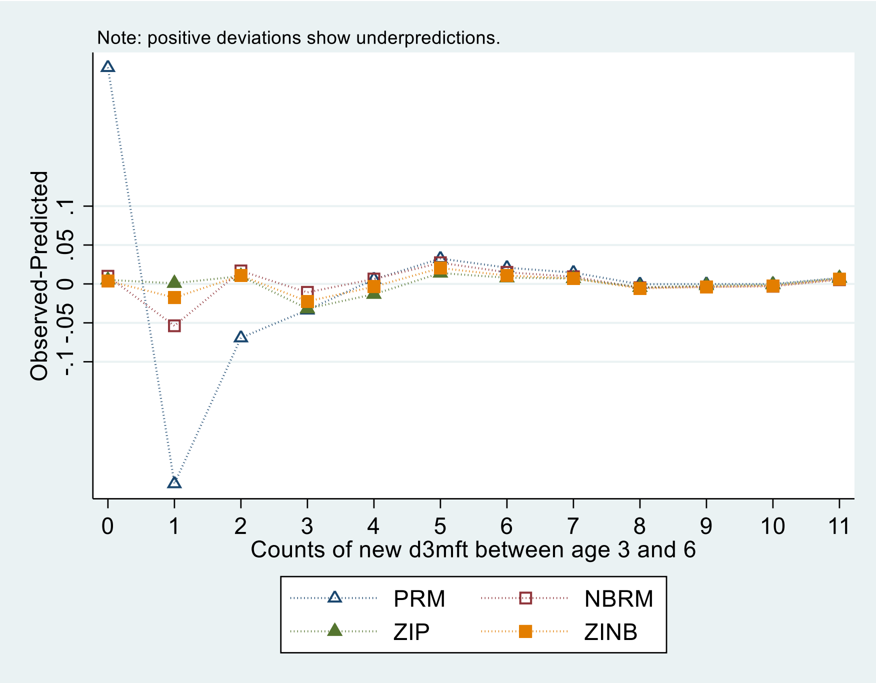
Four different models (Poisson, negative binomial, zero-inflated Poisson, zero-inflated negative binomial) were tested to identify the approach that could best address the issues of overdispersion and excess zeros.

Also, different covariates were tested to account for the latent class of excess zeros (variable that could explain the excess, i.e. account for a fraction of subjects who did not developed caries at 6). The final model was identified by comparing goodness-of-fit statistics (GoF) (pseudo R-squared, AIC, BIC) and evaluating differences between observed and average estimated probabilities for each count (Figure 22).



The red line is a locally weighted regression line (smoothed) that captures the crude relationship between the dependent and independent variables without making assumptions on its shape (non-parametric).

Figure 21: Scatterplot of 'counts of new d3mft between ages 3 and 6 years' by sucrose intake at age 3 years



PRM = Poisson; NBRM = negative binomial; ZIP = zero-inflated Poisson; ZINB = zero-inflated negative binomial. The ZINB model shows the best fit (also from comparison of GoF statistics) as most differences are close to the zero line.

Figure 22: Differences between observed and average estimated probabilities for each count (1 to 11, observed outcome range)

Table 30: Final model on the association between sucrose intake at age 3 years and the outcome 'counts of new d3mft between ages 3 and 6 years'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The IRR for sucrose intake is expressed per 10 g/day increase. The negative binomial distribution is assumed as the estimated dispersion parameter is not equal to zero. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (d3mft at age 3) is not shown (IRR: 0.21, 95% CI: 0.03, 1.61, p = 0.133)

Outcome variable: outC_3_6 (Counts of new d3mft between age 3 and 6), n=128

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 3 yrs - 10 g/d unit				
per unit	1.21	0.18	0.191	(0.91 to 1.61)
Daily tooth-brushing Y/N - 6 yrs				
No*	1			
Yes	0.42	0.25	0.145	(0.13 to 1.35)
D3mft primary - 3 yrs				
per unit	0.64	0.14	0.042	(0.42 to 0.98)

Table 31: Non-linear model on the association between sucrose intake at age 3 years and the outcome 'counts of new d3mft between ages 3 and 6 years'. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. Categorical IRRs are estimated per quartiles of Sucrose intake. The negative binomial distribution is assumed as the estimated dispersion parameter is not equal to zero. The covariate chosen

as potential explanatory variable for the excess of zeros (d3mft at age 3) is not reported (IRR: 0.20, 95% CI: 0.03, 1.51, p = 0.119)

Outcome variable: outC_3_6 (Counts of new d3mft between age 3 and 6), n=128

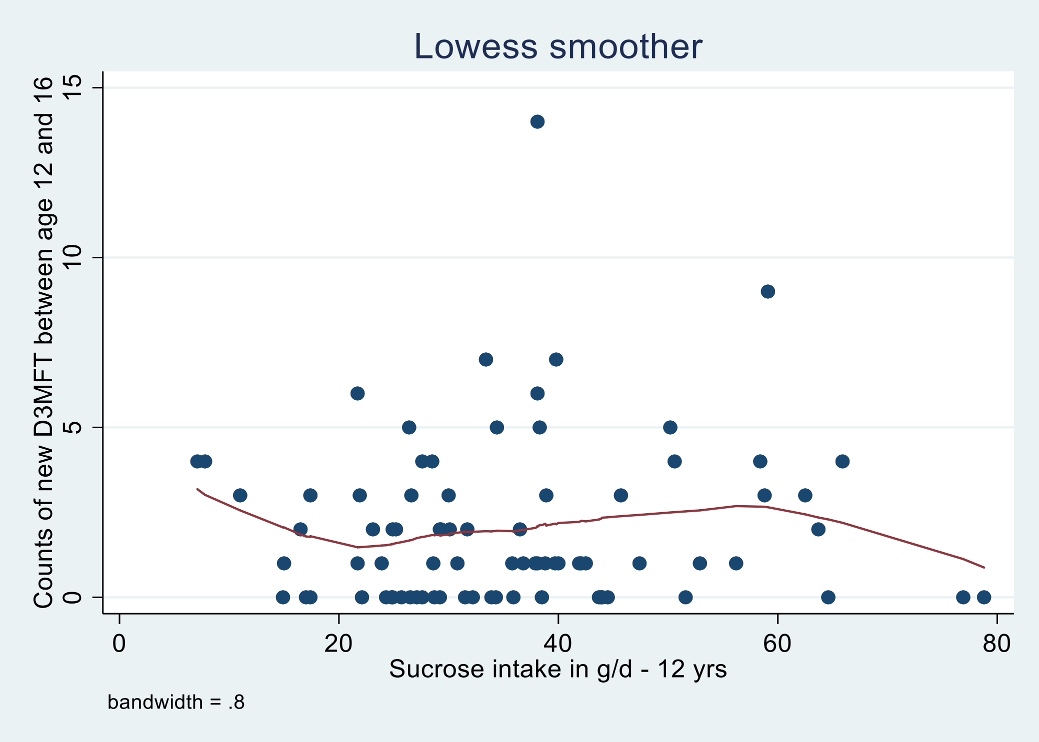
Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 3 yrs - quartiles				
1*	1			
2	0.59	0.37	0.404	(0.17 to 2.05)
3	0.66	0.36	0.445	(0.23 to 1.91)
4	1.54	0.73	0.359	(0.61 to 3.89)
Daily tooth-brushing Y/N - 6 yrs				
No*	1			
Yes	0.47	0.32	0.268	(0.13 to 1.77)
D3mft primary - 3 yrs				
per unit	0.68	0.13	0.052	(0.47 to 1.00)

3.2.4. Association between sucrose intake at age 12 years and increment of dental caries between ages 12 and 16 years (counts)

The association of sucrose intake at age 12 years with the outcome 'counts of new D3MFT between ages 12 and 16 years' was tested first in univariable and then in multivariable models (applying the same modelling strategy as per indicators), including the following potential confounders: intervention group, gender, caries experience at 12 years old, daily tooth-brushing (Y/N) at 12 years old.

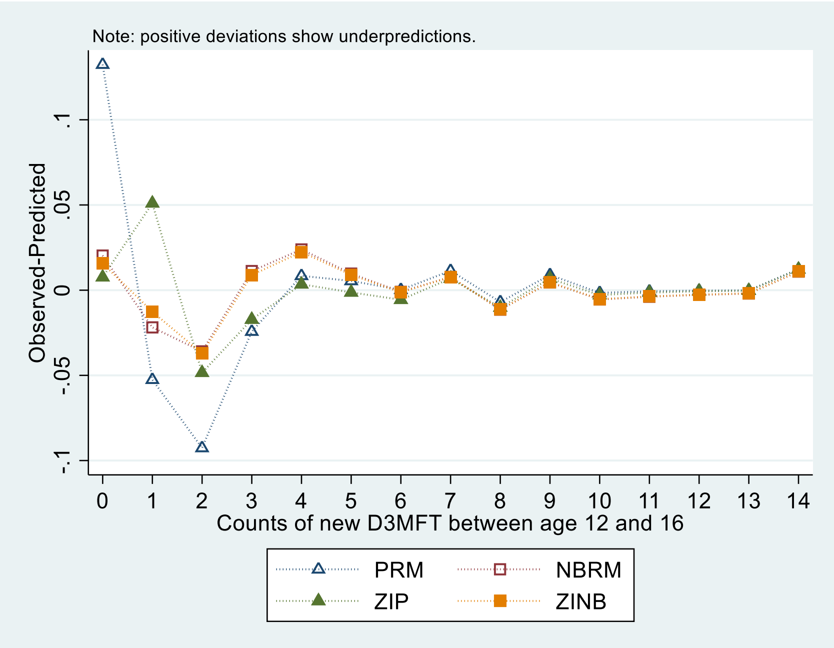
Four different models (Poisson, negative binomial, zero-inflated Poisson, zero-inflated negative binomial) were tested to identify the approach that could best address the issues of overdispersion and excess zeros.

Also, different covariates were tested to account for the latent class of excess zeros (variable that could explain the excess, i.e. account for a fraction of subjects who did not developed caries at 9). The final model was identified by comparing goodness-of-fit statistics (GoF) (pseudo R-squared, AIC, BIC) and evaluating differences between observed and average estimated probabilities for each count (Figure 24).



The red line is a locally weighted regression line (smoothed) that captures the crude relationship between the dependent and independent variables without making assumptions on its shape (non-parametric).

Figure 23: Scatterplot of 'counts of new D3MFT between ages 12 and 16' years by sucrose intake at age 12 years



PRM = Poisson; NBRM = negative binomial; ZIP = zero-inflated Poisson; ZINB = zero-inflated negative binomial. The ZINB model shows the best fit (also from comparison of GoF statistics) as most differences are close to the zero line, however less compared to the previous instance. The NBRM model can be an alternative in this case.

Figure 24: Differences between observed and average estimated probabilities for each count (1 to 11, observed outcome range)

Table 32: Final model on the association between sucrose intake at age 12 years and the outcome 'counts of new D3MFT between ages 12 and 16' years. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The IRR for sucrose intake is expressed per 10 g/day increase. The negative binomial distribution is assumed as the estimated dispersion parameter is not equal to zero. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (D3MFT at age 12 years) is not shown (IRR: 1.16, 95% CI: 0.12, 11.46, p = 0.897)

Outcome variable: outC_12_16 (Counts of new D3MFT between age 12 and 16), n=81

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 12 yrs - 10 g/d unit				
per unit	0.99	0.09	0.942	(0.84 to 1.18)
Daily tooth-brushing Y/N - 12 yrs				
No*	1			
Yes	0.80	0.21	0.397	(0.48 to 1.34)
D3MFT permanent - 12 yrs				
per unit	1.40	0.14	0.001	(1.16 to 1.70)
Intervention group				
1*	1			
2	0.63	0.17	0.085	(0.38 to 1.07)

Table 33: Non-linear model on the association between sucrose intake at age 12 and the outcome 'counts of new D3MFT between ages 12 and 16' years. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. Categorical IRRs are estimated per quartiles of Sucrose intake. The negative binomial distribution is assumed as the estimated dispersion parameter is not equal to zero. The output relative to the covariate chosen as potential explanatory variable for the excess of zeros (D3MFT at age 12) is not shown (IRR: 1.20, 95% CI: 0.22, 6.48, p = 0.828)

Outcome variable: outC_12_16 (Counts of new D3MFT between age 12 and 16), n=81

Covariate	exp(coef.)	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 12 yrs - quartiles				
1*	1			
2	1.35	0.50	0.408	(0.66 to 2.78)
3	1.29	0.42	0.429	(0.69 to 2.42)
4	1.09	0.40	0.818	(0.53 to 2.22)
Daily tooth-brushing Y/N - 12 yrs				
No*	1			
Yes	0.80	0.20	0.386	(0.49 to 1.32)
D3MFT permanent - 12 yrs				
per unit	1.41	0.14	<0.001	(1.16 to 1.70)
Intervention group				
1*	1			
2	0.62	0.16	0.074	(0.37 to 1.05)

Table 34: Negative binomial model on the association between sucrose intake at age 12 and the outcome 'counts of new D3MFT between ages 12 and 16' years. Adjusted IRRs, their standard errors, p-values from Wald tests and 95% CI are reported for each covariate. The IRR for sucrose intake is expressed per 10 g/day increase. The negative binomial distribution is assumed as the estimated dispersion parameter is not equal to zero. Excess of zeros does not seem to be an issue with the distribution of this outcome, so the negative binomial model can be considered an alternative to the zero-inflation mixture model

Outcome variable: outC_12_16 (Counts of new D3MFT between age 12 and 16), n=81

Covariate	IRR	Std. Err.	P>|z|	95% Conf. Interval
Sucrose intake at 12 yrs - 10 g/d unit				
per unit	1.00	0.08	0.983	(0.85 to 1.18)
Daily tooth-brushing Y/N - 12 yrs				
No*	1			
Yes	0.85	0.21	0.519	(0.52 to 1.39)
D3MFT permanent - 12 yrs				
per unit	1.37	0.13	0.001	(1.15 to 1.64)

* Baseline category

References

Chankanka O, Cavanaugh JE, Levy SM, Marshall TA, Warren JJ, Broffitt B and Kolker JL, 2011. Longitudinal associations between children's dental caries and risk factors. Journal of Public Health Dentistry, 71(4), 289–300. doi:10.1111/j.1752-7325.2011.00271.x

Karjalainen S, Söderling E, Sewón L, Lapinleimu H and Simell O, 2001. A prospective study on sucrose consumption, visible plaque and caries in children from 3 to 6 years of age. Community Dentistry and Oral Epidemiology, 29(2), 136–142. doi:10.1034/j.1600-0528.2001.290208.x

Karjalainen S, Tolvanen M, Pienihäkkinen K, Söderling E, Lagström H, Simell O and Niinikoski H, 2015. High sucrose intake at 3 years of age is associated with increased salivary counts of mutans streptococci and lactobacilli, and with increased caries rate from 3 to 16 years of age. Caries Research, 49(2), 125–132. doi:10.1159/000369358

Lewsey JD and Thomson WM, 2004. The utility of the zero-inflated Poisson and zero-inflated negative binomial models: a case study of cross-sectional and longitudinal DMF data examining the effect of socio-economic status. Community Dentistry and Oral Epidemiology, 32(3), 183–189. doi:10.1111/j.1600-0528.2004.00155.x

Preisser JS, Stamm JW, Long DL and Kincade ME, 2012. Review and recommendations for zero-inflated count regression modeling of dental caries indices in epidemiological studies. Caries Research, 46(4), 413–423. doi:10.1159/000338992

Scott Long J and Freese J, 2014. Regression models for categorical dependent variables using Stata (3rd ed.), Stata Press.

Glossary, abbreviations, and acronyms

AIC	Akaike's information criterion
BIC	Bayesian information criterion
CI	Confidence interval
DMFT	Decayed, missing and filled teeth
FJ	Fruit juice
GoF	Goodness-of-fit
IFS	Iowa Fluoride Study
IRR	Incidence rate ratios
NBRM	Negative binomial
NC	Non-cavitated
OR	Odds ratio
PRM	Poisson
SES	Socio-economic status
ZINB	Zero-inflated binomial
ZIP	Zero-inflated poisson